A histidine-rich Pseudomonas metallothionein with a disordered tail displays higher binding capacity for cadmium than zinc

Habjanič, Jelena ; Zerbe, Oliver ; Freisinger, Eva

Abstract: Metallothioneins (MTs) are crucial players in metal-related physiological processes. They are characterized by a high cysteine content and unique metal binding properties resulting in specific metal–thiolate clusters formation. Here we present the first NMR solution structure of a Pseudomonas MT, PflQ2 MT, using the strain P. fluorescens Q2-87. It consists of a metal binding domain and an intrinsically disordered C-terminal tail, that was not observed in other MTs so far. While not influencing the structure or function of the metal binding domain, the tail contains a potential binding motif that might be important in so far undiscovered biological interactions. Unusual is the different metal binding capacity for three ZnIIversus four CdII ions that results in two novel metal-cluster topologies. Nevertheless, the affinity for the fourth CdII ion is reduced due to transient coordination. PflQ2 MT contains an unusually large number of four histidine residues, of which only one is involved in metal ion binding. The three non-coordinating histidine residues influence neither the protein fold nor the stability in vitro. We demonstrate that reinstatement of a second coordinating histidine residue, observed for cyanobacterial MTs, in place of a non-coordinating residue in Pseudomonas MTs, decreases the kinetic lability of the cluster, while preserving the overall metal ion binding stability and the protein fold. Since high thermodynamic stability combined with high kinetic lability of metal binding are mechanistic features critical for the function of MTs, the observed replacement might be advantageous for Pseudomonas MTs with respect to metal ion handling in vivo.

DOI: https://doi.org/10.1039/c8mt00193f

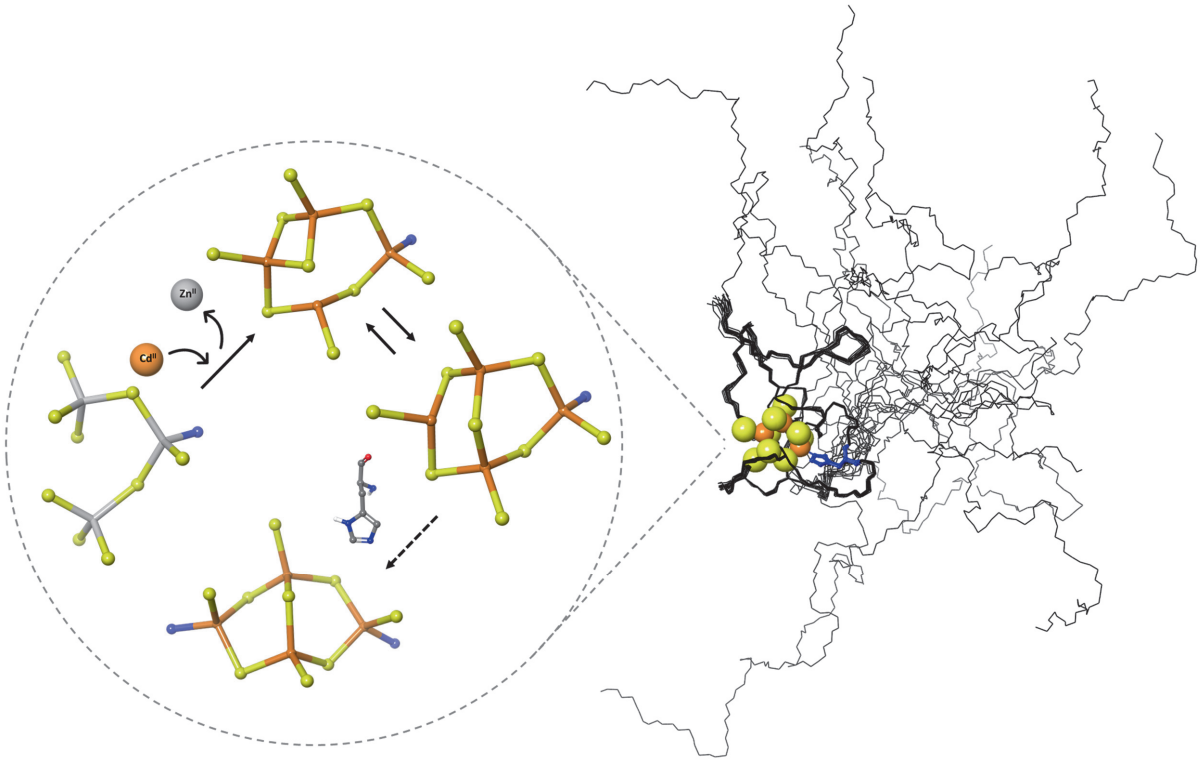
A histidine-rich *Pseudomonas* metallothionein with a disordered tail displays higher binding capacity for cadmium than zinc

Jelena Habjanič, Oliver Zerbe and Eva Freisinger*

Department of Chemistry, University of Zurich, Zurich, Switzerland

E-mail: freisinger@chem.uzh.ch

Dedicated to the memory of Prof. Dr. Silvia Atrian



Abstract

Metallothioneins (MTs) are crucial players in metal-related physiological processes They are characterized by a high cysteine content and unique metal binding properties resulting in specific metal-thiolate clusters formation. Here we present the first NMR solution structure of a *Pseudomonas* MT, PflQ2 MT, using the strain *P. fluorescens* Q2-87. It consists of a metal binding domain and an intrinsically disordered C-terminal tail, that was not observed in other MTs so far. While not influencing

the structure or function of the metal binding domain, the tail contains a potential binding motif that might be important in so far undiscovered biological interactions. Unusual is the different metal binding capacity for three Zn^{II} *versus* four Cd^{II} ions that results in two novel metal-cluster topologies. Nevertheless, the affinity for the fourth Cd^{II} ion is reduced due to transient coordination. PflQ2 MT contains an unusually large number of four histidine residues, of which only one is involved in metal ion binding. The three non-coordinating histidine residues influence neither the protein fold nor the stability *in vitro*. We demonstrate that reinstatement of a second coordinating histidine residue, observed for cyanobacterial MTs, in place of a non-coordinating residue in *Pseudomonas* MTs, decreases the kinetic lability of the cluster, while preserving the overall metal ion binding stability and the protein fold. Since high thermodynamic stability combined with high kinetic lability of metal binding are mechanistic features critical for the function of MTs, the observed replacement might be advantageous for *Pseudomonas* MTs with respect to metal ion handling *in vivo*.

Significance to metallomics

Metallothioneins (MTs) are ubiquitous small cysteine-rich proteins, that obtain their characteristic structure through coordination of metal ions arranged in clusters. PflQ2 MT from *Pseudomonas fluorescens* is closely similar in structure and sequence to bacterial SmtA. However, it lacks one of the two His ligands resulting in reduced kinetic but not thermodynamic metal cluster stability. The absence of typical metal-responsive transcription factors in the respective operon as well as peculiarities such as a disordered Cys-free C-terminal tail and a charged His residue suggest an altered biological role for Pseudomonas when compared to bacterial MTs.

Abbreviations: MT, metallothionein; PflQ2 MT, metallothionein from *Pseudomonas fluorescens* Q2-87; sh_PflQ2 MT, shortened version of metallothionein from *Pseudomonas fluorescens* Q2-87; bacMTs, bacterial metallothioneins; SEC, size-exclusion chromatography; ESI-MS, electrospray ionization mass spectrometry; F-AAS, flame atomic absorption spectroscopy; Tris, tris(hydroxymethyl)aminomethane; TCEP, tris(2-carboxyethyl)phosphine; 5F-BAPTA, 1,2-bis-(2-amino-5-fluorophenoxy)ethane-N,N,N',N'-tetra-acetic acid ; 2-PDS, 2,2’ dithiopyridine; GST, glutathione S-transferase; TEV, tobacco etch virus

Introduction

Metallothioneins (MTs) are a superfamily of cysteine-rich, low molecular weight (5-12 kDa) proteins found in all phyla. They have the ability to coordinate various transition metal ions with a d^{10} electron configuration forming specific metal-thiolate clusters. Depending on the organism, essential (Zn^{II}, Cu^{I}) and/or toxic metal ions (e.g. Cd^{II}, Hg^{II}) can be bound, defining the role of the respective MT in metal homeostasis, protection against metal toxicity, or even against oxidative or nitrosative stress imposed by the redox capabilities of the cysteine thiolate groups.[1-5] Usually, MTs are located in the cytoplasm, but in higher organisms, they are also found in the extracellular space (e.g. blood plasma, cerebrospinal fluid) suggesting that their biological roles might also be localization-dependent.[6]

Although MTs are known since the 1950s, bacterial metallothioneins (bacMTs) were not discovered until 1979 when the first bacMT from the cyanobacterium *Synechococcus elongatus* PCC 7942 was characterized.[7,8] This MT, later named SmtA, is a 56 residue protein, which in addition to nine cysteines also contains three histidine and two tyrosine residues; at that time an unknown feature for a MT. The solution NMR structure of the zinc form of this protein revealed a $Zn_4Cys_9His_2$ cluster, and for the first time demonstrated that histidine residues can be involved in the binding of metal ions in MTs.[9] Regarding its physiological function, SmtA plays a significant role in zinc detoxification and a minor role in cadmium detoxification.[10,11]

A second representative of the bacMT family, MymT, was identified serendipitously roughly 30 years later when the pathogenic Gram-positive bacterium *Mycobacterium tuberculosis* was screened for resistance to a potential drug.[12] MymT is a copper-specific MT with a major role in copper homeostasis and detoxification. Its discovery suggested that bacMTs were probably widely overlooked considering that this MT was not even recognised as a hypothetical protein in the initial genome annotation.[13]

Pseudomonas MTs are an emerging subfamily of bacMTs. Out of the 215 complete *Pseudomonas* genomes sequenced to date, 168 (~80%) list one MT gene, and hence MTs seem indeed to be common in bacteria and likely play a significant role in their viability.[14] Only recently, several studies reported the importance of metal metabolism in *Pseudomonas* and revealed a remarkable degree of adaptation to both limitation and excess of metals.[15-20] While no actual *in* vivo expression of MTs was detected, it

should be considered that the high complexity of a fine-tuned mechanisms can prevent detection of all components, and in the present case might contribute to the overall insufficient knowledge about the metal metabolism of *Pseudomona*s. Interestingly, however, up-regulation of an MT gene was detected in *P. aeruginosa* PAO1 upon formation of biofilms, that make cells more resistant to heavy metal stress than free-swimming cells.[21-24] Since most studies are routinely done with planktonic cells, this might explain why *Pseudomonas* MTs were widely overlooked so far.

Pseudomonas and cyanobacterial MTs possess a vital difference at the genome level: While the cyanobacterial genome contains a well-defined smt operon consisting of the zinc-regulated MT (SmtA) and its divergently transcribed zinc-dependent repressor SmtB separated by a defined operator-promoter region, the *Pseudomonas* MT operon lacks any metal-responsive transcription factor or its common promoter sequence.[25] Moreover, if present at all, the genes immediately upstream and downstream of *Pseudomonas* MTs encode for so far unknown proteins. This indicates that *Pseudomonas* MTs probably evolved to have different functionalities than cyanobacterial MTs and that they might participate in one of many underlying mechanisms of metal homeostasis or have functionalities so far not assigned to MTs.

On the amino acid level, *Pseudomonas* and *Synechococcus* MTs share a conserved Cys pattern, consisting of an N-terminal CxCxxCxC, a central YCC/SxxCAxxH, and a C-terminal Cxxxx(x)CxC motif (Figure S1). The most significant difference, however, is the replacement of one of the two Zn^{II} coordinating histidine residues (H49 in SmtA) with a non-coordinating residue (e.g. Gly, Ala, Ser) in *Pseudomonas* MTs. This might have a significant influence on the metal ion coordination properties. Another difference is the larger number of histidine residues in *Pseudomonas* MTs, especially in *P. fluorescens* strains, that varies between two and eight. As a very peculiar feature, some of the *Pseudomonas* MT sequences have a long Cys-free C-terminal tail.

The focus of our study is to investigate the impact of these features on structural and functional properties using the MT from the *P. fluorescens* strain Q2-87 (PflQ2 MT) as a model protein:

```
  1 NELRCGCPDC HCKVDPERVF NHDGEAYCSQ ACAEQHPNGE PCPAPDCHCE  50
 51 RSGKVGGRDI TNNQLDEALE ETFPASDPIS P                     81
```

In particular, four aspects of *Pseudomonas* MTs will be investigated in detail: the structural relevance of the long C-terminal tail (1), the Zn^{II} and Cd^{II} binding properties (2), the influence of having one coordinating histidine residue less compared to SmtA (3), as well as the function of the additional non-coordinating histidine residues (4).

Materials and methods

Chemicals and solutions

The enzymes and purification systems used for plasmid construction were purchased from Fermentas (Le Mont-sur-Lausanne, Switzerland) and Sigma-Aldrich (Buchs, Switzerland). All buffers and chemicals were ACS grade or comparable from Sigma-Aldrich (Buchs, Switzerland), Merck (Darmstadt, Germany), Chemie Brunschwig (Basel, Switzerland), and Roth AG (Arlesheim, Switzerland). Chelex® 100 resin was from Bio-Rad (Reinach, Switzerland). Isotopically labelled compounds (Tris (D_{11}, 98%), D-glucose (U-$^{13}C_6$, 99%), ammonium chloride (^{15}N, 99%), and cadmium chloride (^{113}Cd, 95.24%)) were purchased from Cambridge Isotope Laboratories (ReseaChem, Burgdorf, Switzerland), while D_2O (99.8%) was purchased from Armar Chemicals (Döttingen, Switzerland). Plasmid pRK793 for the production of tobacco etch virus (TEV) protease in *Escherichia coli* was a gift from David Waugh (Addgene plasmid # 8827).[26] All solutions were prepared using Millipore water and were degassed under vacuum or saturated with argon. Wherever possible, experiments were performed in a Type B Vinyl Anaerobic Chamber (Coy Lab, USA) equipped with a palladium catalyst and operated with a 5% hydrogen/95% nitrogen gas mix.

Protein expression and purification

The coding sequence for the PflQ2 MT (residues 1-81) was optimized for the expression in *Escherichia coli* (Figure S32) and constructed from commercially obtained oligonucleotides (Microsynth AG, Balgach, Switzerland) (Table S2). A Tobacco Etch Virus (TEV) protease cleavage site was introduced at the 5' end of the protein-encoding sequence (Figure S32). Using the upstream BamHI and the downstream Xmal restriction sites, the coding sequence was cloned into the pGEX-4T-1 expression vector, which encodes an N-terminal glutathione S-transferase (GST) tag (Figure S33). The shortened version of the protein lacking the C-terminal tail, sh_PflQ2 MT (residues 1-52), as well as its A44H

mutant were constructed in the same manner. The plasmid encoding for sh_PflQ2 MT was used as a DNA template for the site-directed mutagenesis to construct the H11N, H22N, and H48N mutants.

Proteins were expressed in *E. coli* BL21(DE3). Cells were grown in Luria-Bertani (LB) media containing 100 μg/mL ampicillin. For the preparation of ^{15}N and $^{15}N,^{13}C$ labelled proteins, cells were grown in M9 minimal media supplemented with $^{15}NH_4Cl$ and/or $^{13}C_6$ D-glucose as only nitrogen or carbon sources. After induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an optical density at 600 nm of 0.7, cells were grown for 6 h at 30 °C, harvested by centrifugation, and lysed by sonification. For all chromatographic steps, ÄKTA Pure 25 or ÄKTA Prime Plus systems were used (GE Healthcare, Glattbrugg, Switzerland). The supernatant was applied to a GSTPrep FF 16/10 column (GE Healthcare) pre-equilibrated with 100 mM phosphate buffered saline (PBS buffer) at pH 7.4. The protein was eluted from the column using 10 mM glutathione in 50 mM Tris-HCl, pH 8.0 and dialysed using 1 kDa Spectra/Por 7 Standard Pre-treated Dialysis Tubing (Spectrum Labs, USA) against 10 mM Tris HCl buffer, pH 8.0. For cleavage, the proteins were incubated overnight with TEV protease at room temperature that was prepared as described before with the following modifications.[27] The supernatant was applied on a 5 mL HisTrap HP column (GE Healthcare) pre-equilibrated with IMAC buffer and washed until stabilization of the baseline. TEV protease was eluted using elution buffer and immediately applied on a HiLoad 16/600 Superdex 75pg column (GE Healthcare) pre-equilibrated with elution buffer.

After cleavage with TEV protease, the desired protein was purified from the cleavage mixture again using the HiLoad 16/600 Superdex 75pg column pre-equilibrated with 10 mM ammonium acetate, pH 7.4. Fractions containing the cleaved protein were lyophilized and stored at -20 °C. The overall protein yield was around 25 mg per L of cell culture.

Apo-stocks of the proteins were prepared by resuspending the respective lyophilised protein in 10 mM HCl pH 2.4 to release any bound metal ions, incubated with 100 mM tris(2-carboxyethyl)phosphine (TCEP) for 20 minutes to ensure complete reduction of all Cys residues, and purified by size exclusion chromatography (SEC) using a Superdex Peptide 10/300 GL column (GE Healthcare) pre-equilibrated with 10 mM HCl under anaerobic conditions to separate the buffer components including any potential

metal ions from the apo-protein fraction. The identity of the proteins was confirmed by electrospray ionization mass spectrometry (ESI-MS, Figure S34).

Metal ion coordination studies

Determination of binding capacity

The maximum metal binding capacity was determined by adding an excess of the respective metal ion to the apo-protein and increasing the pH to 7.4. The metal ion excess was subsequently removed by SEC (Superdex Peptide 10/300 GL column, pre-equilibrated with 50 mM Tris HCl, 50 mM NaCl, pH 7.4). Fractions containing the holo-form of the protein were analyzed by flame atomic absorption spectroscopy (F-AAS) on an AA240FS spectrometer (Agilent Technologies AG, Basel, Switzerland) to determine the metal ion concentration and the 2,2' dithiopyridine (2-PDS) assay to determine the protein concentration via thiol quantification.[28] Additionally, each sample was analyzed by ESI-MS.

Zn^{II}/Cd^{II} titrations of apo-MTs

Metal titration experiments were performed in a septum-sealed spectrophotometric cuvette (Hellma AG, Switzerland) by adding increasing amounts of the respective metal ion with a microsyringe (Hamilton, Bonaduz, Switzerland). For each titration, 800 µL of a 10-15 µM protein solution in 10 mM Tris HCl, 50 mM NaCl, pH 7.4 was used. After each titration experiment, the protein was analysed for oxidation by thiolate group quantification with the 2-PDS assay.

Zn^{II}-to-Cd^{II} exchange

To evaluate stoichiometric replacement of Zn^{II} ions by Cd^{II}, the Zn^{II} form of sh_PflQ2 MT in 50 mM Tris-HCl, 50 mM NaCl, pH 7.4, was incubated for ~1 h with four equivalents of Cd^{II}, dialysed against 50 mM ammonium acetate, pH 7.4, overnight at room temperature, and analyzed by ESI-MS.

Metal ion binding stability determination using ^{19}F NMR spectroscopy

Overall apparent stability constants for Zn^{II} and Cd^{II} binding were determined using a modified procedure developed by Hasler et al.[29] 150 µM solutions of the respective metallated protein in 50 mM Tris HCl, 300 mM NaCl, pH 7.4, were incubated with 2.5 mM 1,2-bis-(2-amino-5-fluorophenoxy)ethane-N,N,N',N'-tetra-acetic acid (5F-BAPTA, Biotium, USA) overnight at room

temperature. ^{19}F $\{^{1}H\}$ spectra were acquired using a Bruker AV2-400 spectrometer equipped with a 5 mm QNP probehead at 300 K using a spectral width of 237 ppm, a relaxation delay of 1 s, and 500 scans for each experiment.

Calculations were performed as described, and Zn^{II}/Cd^{II} 5F-BAPTA stability constants were corrected using the program Chelator.[29,30] The initial values for $\log K_{Cd(BAPTA)}$ and $\log K_{Zn(BAPTA)}$ used for the calculations were 11.75 and 9.91, respectively.[31] After correction for the ionic strength (I=0.321 M), temperature (300 K) and pH (7.4) with Chelator the values were 9.96 for $\log K_{Cd(BAPTA)}$ and 8.12 for $\log K_{Zn(BAPTA)}$. Metal ion concentrations were determined by F-AAS, while protein concentrations were determined by the 2-PDS assay.

Zn^{II} ion transfer from MT to apo-alkaline phosphatase

The Zn^{II}-free form of *E. coli* alkaline phosphatase (AP, Sigma Aldrich) was prepared freshly prior to each experiment using Chelex® 100 resin as developed by Csopak.[32] All solutions were pretreated with Chelex® 100 resin to remove any traces of Zn^{II} ions. Zn^{II} transfer from the respective Zn^{II}-MT to apo-alkaline phosphatase was monitored as described.[33] Apo-AP in Tris HCl, pH 8.0, was first mixed with the substrate p-nitrophenyl phosphate (Fluorochem, UK) followed by rapid mixing with the respective Zn^{II}-MT to yield final concentrations of 0.5 μM apo-AP, 5 μM Zn^{II}-MT, and 1 mM p-nitrophenyl phosphate in 10 mM Tris HCl, pH 8.0. The formation of the product p-nitrophenolate was monitored with absorption spectroscopy at 400 nm for 1 h.

Mass spectrometry

ESI-MS measurements were performed on a Synapt G2 quadrupole time-of-flight spectrometer (Waters, UK) with a capillary voltage of 2.8 V, a cone voltage of 40 V and a source temperature of 80 °C. The apo-forms were analyzed in MeOH:0.15% aqueous formic acid (1:1) and the holo-forms in MeOH:30 mM aqueous ammonium acetate (1:1). The m/z data were deconvoluted using the MaxEnt1 software with a resolution of the output mass of 0.1 Da/channel and a Uniform Gaussian Damage Model at the half height of 0.1 Da.

NMR spectroscopy

Sample preparation

Different metallated species were prepared by dropwise addition of an exact amount of the respective metal ion to the apo-protein solution and increasing the pH to 7.4 with 1 M Tris-D_{11} and/or small additions of NaOH. Samples were lyophilized and resuspended in a buffer containing 50 mM Tris-D_{11} HCl, 50 mM NaCl, pH 7.4 (90% H_2O/10% D_2O) prior to the measurements. The average concentration of ^{15}N (^{13}C,^{15}N) labelled samples and those used for ^{113}Cd measurements was around 0.5 mM and 1-2 mM, respectively. Referencing of ^{1}H chemical shifts was done with the external standard 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, 0.2% pH 7.5), while ^{15}N and ^{13}C frequencies were referenced indirectly.

Experiments were recorded on a Bruker 600 MHz spectrometer equipped with a CRYO TCI triple-resonance probe (^{1}H, ^{13}C, ^{15}N) if not stated otherwise.

Assignment procedure

Protein backbone resonances were assigned by picking peaks in [^{15}N,^{1}H]-HSQC, HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, and HBHA(CBCACO)NH spectra and then using standard triple resonance backbone assignment that provides intra- and interresidual C^{α} and C^{β} correlations (CBCA(CO)NH, HNCACB) or C' resonances (HNCO, HN(CA)CO). To assign protein side-chains, H^{α} and H^{β} resonances obtained from a HBHA(CBCACO)NH spectrum were used to obtain H-C^{α} and H-C^{β} correlations. Using these as anchors the remaining side-chain correlations were determined from (H)CCH-TOCSY and the H(C)CH-TOCSY spectra. Sidechain amides of glutamines and asparagines were assigned from the 3D ^{15}N-resolved NOESY spectrum. Aromatic side chains were assigned with the help of (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE. Annotations of protons of the His π system were supported by a long-range [^{15}N,^{1}H]-HSQC spectrum. Due to the high similarity of the imidazole proton resonances of all four histidine residues, the correctness of H-C correlations was additionally confirmed by inspecting spectra of the His-to-Asn mutants. Upper distance restraints were derived from 120 ms ^{15}N- or ^{13}C- (aromatic and aliphatic) ^{13}C-resolved NOESY spectra.

The cluster topology was established from 1D ^{113}Cd and [^{113}Cd,^{1}H]-HSQC-TOCSY spectra recorded on a Bruker 500 MHz spectrometer equipped with a CRYO 5 mm QNP probe in the temperature range from 275 to 320 K. The ^{113}Cd chemical shifts were referenced to an external 0.1 M $Cd(ClO_4)_2$ solution. All spectra were processed using TopSpin (version 3.5) and analysed in the program CARA (version 1.9.1).[34] Backbone chemical shift perturbations between different constructs were calculated using the following equation (1):[35]

$$\Delta\delta_{N,H} = \sqrt{\frac{\Delta\delta_H^2 + \frac{\Delta\delta_N^2}{5}}{2}} \quad (1)$$

Structure calculation

Automatic picking in ^{15}N-, ^{13}C-, and aro-^{13}C-NOESY-HSQC spectra was done by the ATNOS/CANDID algorithm using UNIO´10 (version 2.02.).[36] Structures were calculated with CYANA (version 3.97) using UNIO-derived unassigned NOESY peak lists, and dihedral angle restraints from TALOS-N.[37,38] The 20 lowest energy conformers were refined in explicit water using CHARMM parallhdg5.3 force-field implemented in XPLOR-NIH (version 4.0). The positions of all Cd^{II}/Zn^{II}, S^{γ} and $N^{\varepsilon 2}$ of H36 and additionally $N^{\delta 1}$ of H44 in the A44H mutant were held fixed during the refinement.

To ensure tetrahedral binding of metal ions (Zn^{II}/Cd^{II}) to ligands (Cys/His), four additional metal pseudo-residues were incorporated into the CYANA library: the CDC residue represents the $CdCys_4$ coordination site, the CDH residue $CdCys_3His$, the ZNH residue $ZnCys_3His$, and the ZNC residue the $ZnCys_4$ coordination site.

All metal pseudo-residues were created using the Maestro program (Schrödinger, USA) in a similar manner as described before.[39] They consist of four dummy ligand atoms (L_1-L_4) in a tetrahedral arrangement around the metal centre (M) with M-L_x distances of 2.5 Å for Cd-Cys distances, 2.3 Å for Cd-His, 2.3 Å for Zn-Cys, and 2.0 Å for Zn-His distances. To impose tetrahedral coordination and correct distances, the distances between the dummy ligand atoms (L_x) and the respective ligand atom of the protein (that is sulphur for cysteine and nitrogen for histidine) were restrained to a maximum value of 0.1 Å. At the same time, to make sure that every metal ion is coordinated by four ligands and every ligand is connected to at least one metal ion, an additional upper distance of 0.1 Å was implemented

between all possible connecting atoms (S/N) and dummy ligand atoms (L_x). Structures were first calculated using ambiguous restraints generally allowing all Cys residues to bind to any of the metal ions, except for H36 which was restricted to act as a terminal ligand based on the long-range [^{15}N,^{1}H]-HSQC spectra. Subsequently, experimentally observed connectivities from [^{113}Cd,^{1}H]-HSQC-TOCSY spectra were implemented into the calculation (Table S3). More details on the structure calculation protocol will be reported elsewhere. Details of structure calculation and geometry are summarised in Table S4.

Studying backbone amide dynamics using $^{15}N\{^{1}H\}$-NOE

^{15}N{^{1}H}-NOE experiments were recorded with a steady-state ^{15}N{^{1}H}-NOE version of the ^{15}N-HSQC experiment using an experiment repetition delay of 5 s. Two [^{15}N,^{1}H]-HSQC spectra, one with and one without amide proton saturation, were recorded and the ^{15}N{^{1}H}-NOE values were computed from the peak volume ratios of NOE-to-reference experiments.

Determination of pK_a values of histidine residues

The pK_a was determined from measurements of ^{13}C chemical shifts of His residues in a pH titration. The titration was performed from pH 4.0 to pH 9.5 at 300 K. A 0.3 mM stock solution of Cd_4sh_PflQ2 MT in 50 mM Tris-D_{11} HCl, 50 mM NaCl, was divided into five samples and the pH was varied by addition of small amounts of diluted HCl or NaOH. For each pH value, [^{13}C,^{1}H]-HSQC spectra, centred in the aromatic ^{13}C region, were measured, and the chemical shift of histidine $C^{\varepsilon 1}$ and $H^{\varepsilon 1}$ was plotted against pH. pK_a values were obtained from fitting data against equation (2) (OriginPro2017), where δ_{HA} is the chemical shift at the lowest pH and δ_A at the highest pH.

$$\delta_{obs} = \frac{\delta_{HA} + \delta_A \cdot 10^{pH-pKa}}{1+10^{pH-pKa}} \quad (2)$$

Results and discussion

The metallothionein from *Pseudomonas fluorescens* Q2-87 (PflQ2 MT) has a number of features that are different from other bacterial MTs. These features include the extra, presumably flexible, tail, and

the high number of His residues. In the following, we individually describe in which way these features influence the structure and metal-binding capability of this protein.

The relevance of the C-terminal tail

A distinct feature of PflQ2 MT is the presence of a long Cys- and His-free C-terminal tail that is found only in certain MT sequences from different *Pseudomonas* strains. These C-terminal sequences contain multiple charged residues, often of equal charge positioned next to each other. It is also noteworthy that the sequence of the last eleven amino acid residues (71-81; ETFPASDPISP) is conserved in all *Pseudomonas* MTs that contain this tail (Figure S1). Analysis of the PflQ2 MT sequence with disorder prediction servers results in a high probability of disorder for the last ~30 amino acids (Figure S2).[40] To experimentally confirm this prediction but also to investigate the influence of the tail on the metal binding abilities and on the structure, the full-length protein PflQ2 MT and its shortened version, lacking the tail, (sh_PflQ2 MT) were compared using NMR and optical spectroscopy.

To assess the metal ion binding capacity, the apo-structure of both protein forms were titrated with increasing amounts of Cd^{II} and Zn^{II} and the characteristic thiolate-to-metal charge transfer (LMCT) bands at 230 (Zn^{II}), and 250 nm (Cd^{II}) monitored by UV spectroscopy (Figure 1, Figure S3). For both protein forms, the maximum absorptivity is reached after the addition of four equivalents of Cd^{II} or three equivalents of Zn^{II}, respectively. MS spectra show the presence of Cd_3 and Cd_4 species for both Cd^{II}-saturated protein forms but only the Zn_3 species for the Zn^{II}-loaded forms (Figure S4). Hence, the metal ion binding capacity of both proteins is not influenced by the presence of the tail. To analyse the chemical environment of the coordinated Cd^{II} ions in the two Cd_4 forms in more detail, 1D ^{113}Cd spectra were collected. In both cases, four signals are observed, and they have the same chemical shifts, which corroborates that also the chemical environment of the coordinated metal ions is identical for full-length and shortened protein and hence independent of the presence or absence of the tail (Figure 2). Metal ion binding stabilities of both protein forms with Zn^{II} and for Cd^{II} were determined using the chelator 5F-BAPTA and ^{19}F NMR spectroscopy. The overall apparent stability constants for Cd^{II}-binding yield comparable values for both protein forms ($\log K_{Cd}$ value of 10.93 ± 0.10 for PflQ2 MT and 10.96 ± 0.01

for sh_PflQ2 MT), while the Zn^{II} binding constant for PflQ2 MT is slightly lower ($\log K_{Zn}$ 9.69 ± 0.13) than for sh_PflQ2 MT ($\log K_{Zn}$ 10.31 ± 0.08) (Figure S5).

Finally, the structural impact of the C-terminal tail was analyzed by solving the structure of Cd_4PflQ2 and Cd_4sh_PflQ2 MT by solution NMR. Comparison of the 20 lowest-energy structures of each protein form revealed that the C-terminal tail is not influencing the rest of the protein structure (Figure 3a, Figure S6). The backbone dynamics study for PflQ2 MT (1-52) and sh_PflQ2 MT show only minimal differences in $^{15}N\{^1H\}$-NOE values (Figure 3b). Accordingly, both structures can be regarded as closely similar in terms of rigidity and fold. Therefore, the influence of the C-terminal tail on the overall protein structure appears to be negligible. Chemical shift comparison of the metal binding domain of the two protein forms revealed, however, certain differences in solvent exposed, charged or aromatic residues that might arise from electrostatic interactions with the highly charged tail Figure S7). Backbone dynamics studies show a gradual decrease of $^{15}N\{^1H\}$-NOE values from residue R51 on (Figure 3b), which, however, only level off to -0.5 after residue 76. This indicates that there is still some, albeit rather loose, interaction of the tail with the folded part of the protein, and that it resembles a molten globule (Figure S6).[41]

Thus, the prediction, that the C-terminal tail is highly flexible and even intrinsically disordered was confirmed. Interestingly, the prediction server identified residues 68 to 75 (ALEETFPA) in the full-length PflQ2 MT as a potential binding region (Figure S2):[42,43] As a general feature, binding regions of intrinsically disordered proteins cannot form local interactions between neighbouring residues due to the high numbers of neighboring residues of equal charge, but they are able to form energetically favourable interaction with a globular target protein upon binding. These interactions generally either serve interaction-based functions (e.g. inhibitors, activators, signal carriers, degrons), entropic chain functions (e.g. entropic bristles and clocks) or chaperone functions (e.g. RNA and protein chaperones).[44] Therefore, even though this novel feature of a C-terminal tail is not significantly influencing the structure or metal binding properties of PflQ2 MT, it could potentially have a crucial cellular function. However, since the occurrence of intrinsically disordered regions is unprecedented for MTs, for now, we can only

speculate about its possible roles, and the question why it is present only in specific *Pseudomonas* MT sequences remains unclear.

Due to the structural similarity of PflQ2 MT and sh_PflQ2 MT and the more straightforward assignment of NMR spectra of the shorter form, sh_PflQ2 MT was used for all further investigations described herein.

Differences invoked by Cd^{II} versus Zn^{II} coordination

For a long time, binding of Zn^{II} and Cd^{II} to metallothioneins was considered to be isostructural, despite the higher thiophilicity and the larger ionic radius of Cd^{II} and hence the longer metal-to-ligand bonds lengths compared to Zn^{II}.[45] Accordingly, Cd^{II} is widely used as a probe for the spectroscopically silent Zn^{II}.[6] However, in the last decade, multiple examples of MTs emerged that challenge this assumption by showing a preference either for Cd^{II} or Zn^{II}.[31,39,46,47] In addition, a positive correlation has been proposed between the formation of homogeneous well-folded structures upon binding of a specific metal ion and the biological relevance of that metal.[46-57] Due to the promiscuous nature of MTs, their metal preference is probably also strongly affected by the local abundance of the respective metal ion in the native organism or its environment.[58]

As shown above and in contrast to SmtA, the *Pseudomonas* MT PflQ2 has a lower metal ion binding capacity for Zn^{II} than for Cd^{II}. Nevertheless, to a certain extent, PflQ2 MT seems to be able to interact with a 4th Zn^{II} ion as the [^{15}N,^{1}H]-HSQC spectrum of a solution containing precisely four equivalents of Zn^{II} reveals the presence of more than one species (Figure S8A). The binding affinity of this 4th Zn^{II} ion must be, however, rather low, since in the ESI-MS spectrum only the Zn_3 species is observed (Figure S8B). Accordingly, the Zn_4 species is most likely biologically not relevant as the 4th metal ion should be easily removed by cellular chelators such as glutathione. Another significant difference to SmtA is the absence of a kinetically inert Zn^{II} binding site in PflQ2 MT:[59] all Zn^{II} ions bound to PflQ2 MT can be exchanged by Cd^{II} ions as confirmed by ESI-MS (Figure S9). The absence of Zn^{II} regulatory elements in the PflQ2 MT operon and the reduced binding capacity of PflQ2 MT in combination with the absence of a kinetically inert Zn^{II} binding site suggests that *Pseudomonas* MTs play a reduced role in zinc detoxification in comparison to SmtA.

In Cd_4sh_PflQ2 MT one of the Cd^{II} ions is less tightly bound and can be partially removed during SEC or MS measurements (Figure S4). Mixed Cd_3/Cd_4 samples display twice as many peaks in the [^{15}N,^{1}H]-HSQC spectrum compared to samples of the homogenous Cd_4 species, and also give rise to additional peaks in the 1D ^{113}Cd spectrum (Figure S10). A homogenous Cd_3 species can be prepared by incubating the apo-protein with precisely three equivalents of Cd^{II}. Such a sample shows a single mass peak of Cd_3sh_PflQ2 MT in ESI-MS spectrum and formation of one distinct and well-folded species in the [^{15}N,^{1}H]-HSQC spectrum and corresponding backbone dynamics studies (Figure S11AC, Figure S12A). Experiments with the chelator 5F-BAPTA reveal that the overall Cd^{II} binding constant is by almost one order of magnitude larger for Cd_3sh_PflQ2 MT compared to Cd_4sh_PflQ2 MT ($\log K_{Cd3}$ 11.94 ± 0.09 *versus* $\log K_{Cd4}$ 10.96 ± 0.01). Considering that 5F-BAPTA removes 0.6 equiv. of Cd^{II} from the Cd_3 species and 1.9 equiv. from the Cd_4 species, the lower binding affinity must be attributed to a significantly weaker bound 4th Cd^{II} ion. Since the formation of a well-folded structure is quite unusual for an only partially metallated MT species, extensive NMR studies were performed comparing the Cd_4, Cd_3, and Zn_3 species of sh_PflQ2 MT.

A comparison of the chemical shift differences between the Zn_3 and Cd_3 species reveals that changes occur only for the cysteine residues and certain residues in their close vicinity (Figure S13) and indicates that the structures of the later metalloforms must be very similar. More significant chemical shift perturbations (Figure S14A) are observed between the Zn_3 and Cd_4 species; however, analysis of the respective 20 lowest-energy structures again shows a surprising degree of similarity (Figure 4).

The metal ion coordination environment in Cd_3- versus Cd_4sh_PflQ2 MT:

Much less straightforward than the comparison of the protein scaffolds is the analysis of the actual metal ion binding sites in the Zn_3 *versus* the Cd_4 species. Generally, the coordinating residues in Cd^{II}-loaded MTs can be identified from correlations of the Cys H^N, H^α and H^β protons to the ^{113}Cd ions, provided they are not undetectable due to exchange-broadening. In contrast, the positioning of the spectroscopically silent Zn^{II} ions in the structure can only be deduced from indirect evidence, e.g. when the NOEs define the structure already so well that metals can be added unambiguously. Hence, the most reliable approach was to study the metal-ligand connectivities in the structurally highly similar Cd_3

species instead and compare these to the Cd_4 species. The 1D ^{113}Cd spectrum of Cd_3sh_PflQ2 MT shows three peaks representing three different Cd^{II} binding sites. The chemical shifts of two ^{113}Cd peaks are well in the range of typical Cys_4 polynuclear coordination sites (695 ppm and 676 ppm) while the third peak at 562 ppm is shifted to a lower value (Figure S11B).[60] Due to the high sensitivity of ^{113}Cd chemical shifts to the nature of the coordination environment, a shift to lower ppm values can indicate an exchange of thiolate by nitrogen (e.g. His) or oxygen ligands (e.g. carboxylate, H_2O) or a deviation from ideal tetrahedral metal binding geometry.[61] To examine if the lower value of the third ^{113}Cd peak is due to the coordination of His residues, a long-range [^{15}N,^{1}H]-HSQC spectrum was recorded, which indeed revealed a ^{113}Cd scalar coupling to $N^{\varepsilon 2}$ of H36 (Figure S11D).

As depicted in Figure 2, the 1D ^{113}Cd spectrum of Cd_4sh_PflQ2 MT shows four peaks representing the four different Cd^{II} binding sites. The chemical shifts of two peaks (713.8 ppm and 615.4 ppm) are again in the range typical for Cys_4 polynuclear coordination, although the peak at 615.4 ppm experienced a pronounced upfield shift compared to the values observed in the 1D ^{113}Cd spectrum of Cd_3sh_PflQ2 MT and is now in the lower ppm range assigned to this coordination environment. The chemical shifts of the two other peaks (585.8 ppm and 535 ppm) are again indicative for an exchange of thiolate ligands by more shielding ligands or a deviation of the coordination geometry (see above). A long-range [^{15}N,^{1}H]-HSQC spectrum shows again ^{113}Cd scalar coupling to $N^{\varepsilon 2}$ of H36, while the other three His residues are non-coordinating (see below). Since coordination of H36 to the Cd^{II} site at 585.8 ppm results in significant violations of NMR distance restrains, the site with Cys_3His coordination is assigned to the peak at 535 ppm. Why the peak at 585.8 ppm has a rather low chemical shift value can be only speculated on: UV spectroscopy showed that addition of the 4th Cd^{II} ion only results in a small increase of the LMCT band at 250 nm of approximately 1500 M^{-1} cm^{-1} (Figure 1A, Figure S3), indicating that coordination of the first three Cd^{II} ions already requires all nine Cys residues as ligands. Accordingly, coordination of the fourth Cd^{II} ion seems to require only a rearrangement of the Cys residues and some of them need to switch from terminal to bridging binding modes (Figure S15). If in addition a deviation from ideal $CdCys_4$ coordination geometry or the involvement of H_2O as an additional ligand is

considered (see below) this would provide an explanation for the second upfield shifted $^{113}Cd^{II}$ peak in the 1D ^{113}Cd spectrum.

The topology of the metal clusters

Metal ion coordination in form of specific metal-thiolate clusters is a conspicuous feature of all MTs. The two most common cluster types with divalent metal ions (Zn^{II}/Cd^{II}) are the β-cluster ($M^{II}_3Cys_9$) found in vertebrate and crustacean MTs and the α-cluster ($M^{II}_4Cys_{11}$) of vertebrate and echinoderm MTs.[62,63] The metal ions in both clusters adopt tetrahedral tetrathiolate coordination geometry. In the β-cluster, the metal ions are connected by three bridging cysteine thiolate ligands to form a six-membered ring, and the remaining six cysteine thiolate groups serve as terminal ligands (Figure 5a). The *α*-cluster has an adamantane-like topology of two fused six-membered rings requiring five bridging and six terminal Cys thiolate ligands (Figure 6b). A particular form of the *α*-cluster is found in SmtA, in which two terminal Cys residues are replaced by two His residues.[9]

Pseudomonas MTs share the same Cys distribution pattern with SmtA, suggesting a similar metal ion-to-thiolate connectivity pattern (Figure S1). However, in all of them, one of the two coordinating His residues observed in SmtA (H49) is replaced by a non-coordinating residue (e.g. A44 in PflQ2 MT). Despite the presence of three other His residues in PflQ2 MT, none of them takes the part of the missing His residue, both in the Cd_3 and the Cd_4 species. Our experiments suggest an M_3Cys_9His cluster stoichiometry for the Zn_3 and Cd_3 species and thus point to an unprecedented cluster type with one additional 10th ligand compared to the typical β-cluster. In addition to the UV titration experiments, binding of all nine Cys residues was also confirmed by the chemical shifts of the Cys-C^{β} in the [^{13}C,^{1}H]-HSQC spectrum, which are all in the range from 29 ppm to 33 ppm typical for deprotonated thiol groups (Figure S16).[64,65] The NMR structure calculations for the Zn_3 species revealed that three Zn^{II} ions are bridged by two Cys-thiolate groups, and H36 is coordinating in terminal mode to the central Zn^{II} ion (site D in Figure 5a). We suggest that the 10th ligand replaces a bridging Cys in the α-cluster topology, thereby creating two additional terminal coordinating thiolates and opening the six-membered ring (Figure 5a). The resulting cluster structure satisfies the metal-ligand connectivities observed in the [^{113}Cd,^{1}H]-HSQC-TOCSY of $^{113}Cd_3$sh_PflQ2 MT (Figure S17). Based on this spectrum, the ^{113}Cd

resonance at 676 ppm (Figure S14B) can be assigned to Cd^{II} in site B, while the resonance at 695 ppm results from Cd^{II} in site A.

All data for the Cd_4 species reveal that as in the Cd_3 species all Cys and one His ligand are again involved in the coordination of the metal ions. The resulting M_4Cys_9His stoichiometry features one ligand less than needed for the typical α-type cluster. Accordingly, the cluster could, in theory, adopt a true adamantane topology by converting one of the terminal thiolates groups into a bridging ligand (in Figure 6 this corresponds to converting the Cys residue C7 in part A, right side, into a bridging Cys to metal site C resulting in the structure depicted in part C). This arrangement, however, requires a distance contraction between the two newly bridged metal ions (sites C and D) by approx. 2 Å, which results in significant violations of the experimental distance restraints obtained from the NMR data of Cd_4PflQ2 MT. Instead, the structure calculation using experimentally obtained connectivities (Figure S18) returns a scenario that still resembles a typical α-cluster topology, but with the distinct difference that the Cys ligand 10 (C10) is either bound to site A or C, leaving the respective other site coordinatively unsaturated, i.e. in the form of a $CdCys_3$ site (Figure 6a). In the computed structure, the thiolate S^{γ} of C10 has a distance of 2.50 ± 0.04 Å to the Cd^{II} at site A (Cd(A), 713.8 ppm), 2.45 ± 0.02 Å to Cd(B) (615.4 ppm), and 2.55 ± 0.04 Å to Cd(C) (585.8 ppm). No conformational changes are required to allow the ligand exchange between sites A and C, while C10 remains coordinated to Cd(B). This transient binding mode of C10 to sites A and C can compensate for the missing thiolate ligand compared to the typical α-cluster arrangement. Since coordination of C10 is only experimentally confirmed for site B, and site C has the highest ^{113}Cd upfield shift of the three centers (585.5 ppm) in 1D ^{113}Cd NMR spectra whereas site A (713.8 ppm) is in the range of typical polynuclear Cys_4 coordination, the equilibrium of interchanging topologies depicted in Figure 6a is most likely shifted towards the right. Due to the high solvent accessibility of site C, the partial involvement of H_2O as a fourth ligand also has to be considered.[66,67] Similar ^{113}Cd chemical shifts that arise from increased shielding were also reported in Littorina litorea MT (516 ppm) and the α-domain of human MT-3 (595 ppm) where they were attributed to transient binding or discrepancy in coordination environment, respectively.[39,68] The observed reduced thermodynamic stability of the fourth Cd^{II} ion in the Cd_4sh_PflQ2 MT, and hence Cd(C), corroborates

this interpretation. The proposed equilibrium and thus the interchange of thiolate ligand C10 must be reflected by remarkable intra-cluster dynamics.[69,70] Previous work using ^{113}Cd saturation transfer experiments indeed demonstrated that Cd^{II}-thiolate bonds, while thermodynamically stable, nevertheless may be kinetically labile, and may allow altering the coordination mode very rapidly.[71,72] In that context, it is noteworthy that the ^{113}Cd signal at 535 ppm, corresponding to Cd(D), is much broader compared to the other Cd resonances and only visible at higher temperatures (320 K) while the other three sites remain unaffected. This kind of behaviour indicates that this metal site is on a different dynamic scale due to slower breakage of the thiolate bonds than for the other three metal sites, which is corroborated by the proposal that the latter three share a single thiolate ligand.

As shown above, PflQ2 MT cannot form a stable Zn_4 cluster. Based on the suggested topology of the Cd_4 cluster there are two likely reasons for this, sterics and kinetics. A typical Zn^{II}-thiolate bond is with on average 2.3 Å, and thus shorter than a Cd^{II}-thiolate bond (2.5 Å). This is usually reflected in a 20% decreased cluster size for the Zn^{II} species and hence increases the likelihood of steric repulsion between protein residues.[73] Transient coordination requires the switching of ligand C10 between metal sites A and C, which is more likely due to the higher the kinetic lability of the metal-thiolate bond. As Cd^{II}-thiolate bonds are kinetically more labile than Zn^{II}-thiolate bonds, also the kinetic aspect supports the formation of the Cd_4 cluster.[74-76]

Based on the structure of the M_3Cys_9His three-metal cluster, incorporation of the fourth metal ion in the Cd^{II} species is achieved by converting the two terminal thiolate ligands of C28 and C42 to bridging ones. This, however, creates a frustrated system with an insufficient number of Cys residues for the stable tetrahedral coordination of all metal ions. As described in the previous paragraph transient coordination allows sharing of the bridging C10 ligand between three metal centres. This scenario is supported by the location of the most significant chemical shift changes of the Cys-C^{β} atoms of C12, C28, C42, and C49 upon coordination of the 4th Cd^{II} ion (Figure S12D). It is also noteworthy that despite multiple ligand rearrangements, the metal-ligand connectivities of site B (and D) are preserved during the transition.

Role of non-binding His-residues

For a long time, it was believed that MTs, by definition, neither contain histidine nor any other aromatic amino acid, but with the discovery of bacterial MTs, this paradigm was revised.[9,57] His residues have been proposed to play a role in the stabilisation of metal clusters by reducing the overall charge, by reducing flexibility through H-bond formation, and by lowering the possibility of disulphide bond formation while at the same time maintaining high affinity towards metal ions, in particular Zn^{II}.[77] Apart from SmtA, there is only one other MT structure available that displays metal-coordination via His, the Zn_6E_c-1 MT from *Triticum aestivum*, which contains an isolated, Zn-finger-like $ZnCys_2His_2$ site.[57]

Out of the four His residues in PflQ2 MT only H36, present as $N^{\delta 1}$-H tautomer, is involved in metal ion binding. The non-coordinating H11 and H22 residues occur as $N^{\varepsilon 2}$-H tautomers, while, surprisingly, H48 resides in the doubly-protonated state (Figure 7). The structures of Zn_3- and Cd_4sh_PflQ2 MT show that all non-coordinating His residues are solvent exposed. Some calculated conformers of Cd_4sh_PflQ2 MT suggest that $H^{\varepsilon 2}$ of H22 forms a hydrogen bond with $O^{\varepsilon 2}$ of E34, indicating that H22 might be important for the stabilization of the protein fold. To study the influence of these non-coordinating His residues on the structure and metal binding properties, His-to-Asn mutants were generated for each of them. Asn was chosen to preserve the polar character of the amino acid while removing its positive charge.

All generated mutants (H11N, H22N, H48N) can be readily produced in *E. coli* and purified. No significant changes in metal binding capacity or the overall apparent Zn^{II} and Cd^{II} stability constants were observed (Figure S5, Figure S19). The 1D ^{113}Cd spectra of the respective Cd_4 species reveal only minor shifts of ^{113}Cd peaks (Figure S20) that are probably due to the vicinity of the mutations to coordinating Cys residues. Significant backbone chemical shift perturbations are limited to residues close to the mutation sites (Figure S21-26). Additional significant changes are detected around residue E34 (Figure S22) in the H22N mutant, most likely due to the missing H-bond to H22 or the lack of a ring current effect. However, overall the role of the H22 residue in the stabilization of the protein fold *in vitro* seems to be only minor.

Since the occurrence of a charged histidine residue at physiological conditions (pH 7.4) is unprecedented for MTs, the apparent pK_a values of the histidine rings were determined by monitoring chemical shift changes of $H^{\varepsilon 1}$ and $C^{\varepsilon 1}$ in the pH range from 4.0 to 9.5. The resulting average pK_a values are 6.57 ± 0 .02 for H11, 6.37 ± 0.04 for H22, and 7.34 ± 0.04 for H48 (Figure 8, Figure S24, Table S3) revealing that the pK_a of H48 is shifted by almost one unit and hence 50% of the H48 residues are in the charged state at the pH chosen in our NMR study (pH 7.4). Charged histidines are usually found buried in the catalytic pockets of enzymes and stabilised by hydrogen bonding, but H48 is clearly solvent exposed. In PflQ2 MT, the unusual positive charge of H48 might be stabilized by the two flanking negatively charged Glu residues as observed in other proteins as well.[78] It is noteworthy that we also observe charged histidine residues at physiological pH in two more MTs from *P. fluorescens* Pf0-1 and *P. putida* KT2440 that have seven and four histidines, respectively (Habjanic J., unpublished results). Its preservation, therefore, might indicate a specific functional and/or biological relevance.

Introducing a second coordinating histidine – the A44H mutant

While the amino acid sequence of PflQ2 MT contains several unique features, the Cys distribution pattern is the same as in SmtA (*vide supra*) and consequently also the overall fold of the two proteins is highly similar (Figure S25). In particular, the characteristic zinc finger-like fold observed in SmtA is preserved, which is not unexpected as its two stabilizing residues Y31 and A37 are also conserved in PflQ2 MT. However, PflQ2 MT does not contain a kinetically inert Zn^{II} site and lacks one of the two coordinating His ligands. The impact of this missing ligand on the structure and metal binding properties was investigated using the A44H-sh_PflQ2 MT mutant. Also, this mutant can be prepared without a decrease in expression yield or stability during the purification process. The titration of apo-A44H-sh_PflQ2 with Zn^{II} does not reveal differences in the absorptivity of the LMCT bands at 230 nm compared to the wild-type (Figure S26A), but the determination of the metal ion content with AAS shows that the maximum Zn^{II} binding capacity increased from three to four equivalents (i.e. 3.83 ± 0.15). ESI-MS shows a low-intensity signal for the Zn_4 and a high-intensity signal for the Zn_3A44H-sh_PflQ2 suggesting that the binding strength of the 4th Zn^{II} ion is still significantly lower than for the other three metal ions (Figure 9a). The chemical shift dispersion in the [^{15}N,^{1}H]-HSQC spectrum of Zn_4A44H-

sh_PflQ2 indicates a well-folded protein and is comparable to the spectrum of Cd_4A44H-sh_PflQ2 (Figure S26B).

Like the wild-type protein, the A44H mutant coordinates four Cd^{II} ions based on analytical results and ESI-MS (Figure S27A). However, the UV spectra of the titration of the apo-protein with Cd^{II} ions reveal a distinct difference: the maximum absorptivity at thiol-to-cadmium charge transfer band at 250 nm is lower than for the wild-type protein, and it is already reached upon addition of the 3rd equivalent of Cd^{II} (Figure S27B). This indicates that all Cys residues are already involved in the coordination of the first three Cd^{II} ions and that, in contrast to the wild-type protein, no additional Cys-Cd bonds are formed when the 4th Cd^{II} ion is added. In line with this observation, two experiments strongly suggest that the new residue H44 is involved in the binding of the 4th Cd^{II} ion (but not of the first three): (i) the 1D ^{113}Cd spectra of Cd_3A44H-sh_PflQ2 and Cd_3sh_PflQ2 MT are virtually identical (Figure S28); (ii) cross-peaks from the $N^{\delta 1}$ signal of H44 in the long-range [^{15}N,^{1}H]-HSQC spectrum of $^{113}Cd_4$A44H-sh_PflQ2 MT display scalar coupling to ^{113}Cd (Figure 9b), while in$^{113}Cd_3$A44H-sh_PflQ2 MT H44 is not coordinating (Figure S11D).

The 1D ^{113}Cd spectrum of Cd_4A44H-sh_PflQ2 MT at 320 K differs from the corresponding spectrum of Cd_4sh_PflQ2 MT. The three highest-intensity ^{113}Cd peaks at 719.2, 623.7, and 591.6 ppm are broader and slightly shifted, while the 4th peak at 535 ppm is not observed in the A44H mutant. At lower temperature (275 K) further line-broadening and the appearance of small additional signals at 720 and 631 ppm (and probably even at 605 ppm) is observed (Figure S28). This indicates that the Cd^{II} sites in Cd_4A44H-sh_PflQ2 MT undergo slower exchange processes than in the wild-type protein, which also explains the disappearance of the fourth ^{113}Cd peak that was already broad in the case of the wild-type protein.

The NMR structure of Cd_4A44H-sh_PflQ2 MT (Figure 10a) reveals that the additional binding of the H44 ligand does not cause any significant changes to the protein conformation. Chemical shift changes occur only in the vicinity of the mutation (Figure S29), and no significant changes in backbone dynamics are observed Figure S29A). Due to broadening of cadmium resonances only a limited number of metal-ligand connectivities can be determined in the [^{113}Cd,^{1}H]-HSQC-TOCSY spectrum (Figure S30), and

therefore most metal-ligand connectivities in the calculation input were left ambiguous. Nevertheless, the calculations returned just a single cluster topology, in which the newly introduced H44 residue acts as a terminal ligand at the cadmium site C. Consequently, C10 is now only bridging metal sites A and B (Figure 10b). Hence, the additional ligand allows the formation of a $Cd_4Cys_9His_2$ cluster with high similarity to the one observed in SmtA.[9] It is noteworthy that the coupling of H44 $N^{\delta1}$ to ^{113}Cd, observed in the long-range [^{15}N,^{1}H]-HSQC spectrum, is much smaller ($^{1}J(^{15}N,^{113}Cd)$ =~97 Hz) than the one for H36 $N^{\varepsilon2}$ ($^{1}J(^{15}N,^{113}Cd)$=~142 Hz). Additionally, H36 is stabilized by a strong hydrogen bond connecting $H^{N\delta1}$ of H36 with O of E40 while H44 is solvent exposed. This observation contributes to the assignment of the ^{113}Cd signal at 593.2 ppm to site C since a similar shift (596 ppm) has been reported for metal site C in SmtA.[9]

Surprisingly, introducing the additional histidine has no stabilizing effect on the overall metal ion binding affinity for the fourth Cd^{II} (10.96 ± 0.01 vs 10.78 ± 0.09), while even a slight decrease is observed for the Zn_3 and Cd_3 species (Figure S5). Zn^{II} transfer to apo-alkaline phosphatase occurs with comparable kinetics for Zn_4A44H-sh_PflQ2 and Zn_3sh_PflQ2 MT, while no transfer is observed for Zn_3A44H-sh_PflQ2 MT (Figure S31). Therefore one could hypothesize that the additional His residue plays a role in the modulation of the reactivity of the MT by decreasing its kinetic lability.

Conclusions

The NMR structure of PflQ2 MT represents the second structure of a bacterial MT and the first of a *Pseudomonas* MT. The disordered C-terminal tail has never been observed in other MTs so far. We could demonstrate that this tail has no influence on the metal binding properties of the *Pseudomonas* MT and does not change the overall protein fold. However, this unique feature possesses a potential binding domain for other proteins, but without precise knowledge of the involved physiological processes, this remains highly speculative. The metal binding studies of PflQ2 MT revealed a decreased binding capacity for Zn^{II} (three Zn^{II} ions bound) as compared to Cd^{II} (four Cd^{II} ions bound). By a combination of different spectroscopic methods and determined solution NMR structure, two novel topologies were proposed for the M_3 and M_4 clusters in PflQ2 MT that differ from the typical α- and β-

clusters usually observed in vertebrate and crustacean MTs. The topology of the novel M_3Cys_9His cluster is characterized by an open six-membered ring, while that of the M_4Cys_9His cluster resembles the typical α-type cluster more closely, but with the three metal centres sharing a common fourth ligand thereby compensating for the missing eleventh ligand. Despite the differences in metal ion coordination, the overall protein fold of the Zn_3/Cd_3 and Cd_4 species is highly similar. Although four histidine residues are present, only H36 is coordinating, while H48 shows an unusually high pK_a value and at physiological pH is 50% in the charged, doubly protonated state. We also confirmed conservation of at least one charged histidine residue in other *Pseudomonas* MTs implying a potential relevance. Remarkably, despite its unique metal binding features and metal ion coordination by just one instead of two histidine residues as seen in cyanobacterial MTs, the fold of the *Pseudomonas* MT resembles that of the cyanobacterial MT SmtA. Introducing another histidine at the position of non-coordinating residue in *Pseudomonas* MT, as found in SmtA, does not significantly affect the overall metal binding capability of PflQ2 MT but alters the kinetics of the metal transfer. To conclude, exchange of one out of two binding histidine residues in SmtA (H49) by a non-coordinating residue in PflQ2 MT (A44) increases the kinetic lability of the cluster but has little influence on the thermodynamics, the metal binding capacity, or the protein fold. Taken together, this work will help to initiate more detailed studies geared towards deciphering the function of *Pseudomonas* MTs *in vivo* and their relevance in the viability of the bacteria.

Accession numbers

Chemical shifts, coordinates, and structure factors are deposited in the BMRB and PDB databases: Cd_4PflQ2MT (PDB ID: 6GRV; BMRB ID: 34282), Cd_4sh_PflQ2MT (6GV6; 34288), Zn_3sh_PflQ2MT (6GW8; 34293), and Cd_4A44H-sh_PflQ2MT (6GV7; 34289).

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We thank Dr. Serge Chesnov (Functional Genomic Centre Zurich) for the mass spectra measurements, Simon Jurt (Department of Chemistry, UZH) for the help with refinement of NMR structures, and Dr. Kerstin Möhle (Department of Chemistry, UZH) for help in creating Figure 6C. Special thanks also to Prof. em. Dr. Milan Vašák for continuous discussions. The work was financially supported by the Forschungskredit of the University of Zurich, grant no. [FK-15-085] and [FK-17-091] to JH.

References

1 M. Brouwer, in *Mechanisms of metallocenter assembly*, eds. R. P. Hausinger, G. L. Eichhorn and L. G. Marzilli, VCH, New York, 1996, pp. 235-260.

2 S. R. Stürzenbaum, C. Winters, M. Galay, A. J. Morgan and P. Kille, Metal ion trafficking in earthworms - Identification of a cadmium-specific metallothionein, *J. Biol. Chem.*, 2001, **276**, 34013-34018.

3 W. Feng, F. W. Benz, J. Cai, W. M. Pierce and Y. J. Kang, Metallothionein Disulfides Are Present in Metallothionein-overexpressing Transgenic Mouse Heart and Increase under Conditions of Oxidative Stress, *J. Biol. Chem.*, 2006, **281**, 681-687.

4 D. U. Spahl, D. Berendji-Grün, C. V. Suschek, V. Kolb-Bachofen and K. D. Kröncke, Regulation of zinc homeostasis by inducible NO synthase-derived NO: Nuclear translocation and intranuclear metallothionein Zn^{2+} release, *Proc. Natl. Acad. Soc.*, 2003, **100**, 13952-13957.

5 S. Zeitoun-Ghandour, O. I. Leszczyszyn, C. A. Blindauer, F. M. Geier, J. G. Bundy and S. R. Stürzenbaum, *C. elegans* metallothioneins: Response to and defence against ROS toxicity, *Mol. Biosyst.*, 2011, **7**, 2397-2406.

6 E. Freisinger and M. Vašák, Cadmium in metallothioneins, *Met. Ions Life Sci.*, 2013, **11**, 339-371.

7 M. Margoshes and B. L. Vallee, A cadmium protein from equine kidney cortex, *J. Am. Chem. Soc.*, 1957, **79**, 4813-4814.

8 R. W. Olafson, K. Abel and R. G. Sim, Prokaryotic metallothionein - Preliminary characterization of a blue-green alga heavy metal-binding protein, *Biochem. Biophys. Res. Commun.*, 1979, **89**, 36-43.

9 C. A. Blindauer, M. D. Harrison, J. A. Parkinson, A. K. Robinson, J. S. Cavet, N. J. Robinson and P. J. Sadler, A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 9593-9598.

10 J. W. Huckle, A. P. Morby, J. S. Turner and N. J. Robinson, Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions, *Mol. Microbiol.*, 1993, **7**, 177-187.

11 R. W. Olafson, W. D. McCubbin and C. M. Kay, Primary- and secondary-structural analysis of a unique prokaryotic metallothionein from a *Synechococcus* sp. cyanobacterium, *Biochem. J.*, 1988, **251**, 691-699.

12 B. Gold, H. Deng, R. Bryk, D. Vargas, D. Eliezer, J. Roberts, X. Jiang and C. Nathan, Identification of a copper-binding metallothionein in pathogenic mycobacteria, *Nat. Chem. Biol.*, 2008, **4**, 609-616.

13 C. A. Blindauer, Bacterial metallothioneins: Past, present, and questions for the future, *J. Biol. Inorg. Chem.*, 2011, **16**, 1011-1024.

14 G. L. Winsor, E. J. Griffiths, R. Lo, B. K. Dhillon, J. A. Shay and F. S. L. Brinkman, Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database, *Nucleic Acids Res.*, 2016, **44**, D646-D653.

15 M. D'Orazio, M. C. Mastropasqua, M. Cerasi, F. Pacello, A. Consalvo, B. Chirullo, B. Mortensen, E. P. Skaar, D. Ciavardelli, P. Pasquali and A. Battistoni, The capability of *Pseudomonas aeruginosa* to recruit zinc under conditions of limited metal availability is affected by inactivation of the ZnuABC transporter, *Metallomics*, 2015, **7**, 1023-1035.

16 L. Izrael-Zivkovic, M. Rikalovic, G. Gojgic-Cvijovic, S. Kazazic, M. Vrvic, I. Brceski, V. Beskoski, B. Loncarevic, K. Gopcevic and I. Karadzic, Cadmium specific proteomic responses of a highly resistant *Pseudomonas aeruginosa* san ai, *RSC Adv.*, 2018, **8**, 10549-10560.

17 M. C. Mastropasqua, M. D'Orazio, M. Cerasi, F. Pacello, A. Gismondi, A. Canini, L. Canuti, A. Consalvo, D. Ciavardelli, B. Chirullo, P. Pasquali and A. Battistoni, Growth of *Pseudomonas aeruginosa* in zinc poor environments is promoted by a nicotianamine-related metallophore, *Mol. Microbiol.*, 2017, **106**, 543-561.

18 K. Mumm, K. Ainsaar, S. Kasvandik, T. Tenson and R. Horak, Responses of *Pseudomonas putida* to zinc excess determined at the proteome level: Pathways dependent and independent of CoIRS, *J. Proteome Res.*, 2016, **15**, 4349-4368.

19 V. G. Pederick, B. A. Eijkelkamp, S. L. Begg, M. P. Ween, L. J. McAllister, J. C. Paton and C. A. McDevitt, ZnuA and zinc homeostasis in *Pseudomonas aeruginosa*, *Sci. Rep.*, 2015, **5**, 13139.

20 J. Quintana, L. Novoa-Aponte and J. M. Arguello, Copper homeostasis networks in the bacterium *Pseudomonas aeruginosa*, *J. Biol. Chem.*, 2017, **292**, 15691-15704.

21 M. Hentzer, L. Eberl and M. Givskov, Transcriptome analysis of *Pseudomonas aeruginosa* biofilm development: anaerobic respiration and iron limitation, *Biofilms*, 2005, **2**, 37-61.

22 G. M. Teitzel and M. R. Parsek, Heavy metal resistance of biofilm and planktonic *Pseudomonas aeruginosa*, *Appl. Environ. Microbiol.*, 2003, **69**, 2313-2320.

23 R. D. Waite, A. Paccanaro, A. Papakonstantinopoulou, J. M. Hurst, M. Saqi, E. Littler and M. A. Curtis, Clustering of *Pseudomonas aeruginosa* transcriptomes from planktonic cultures, developing and mature biofilms reveals distinct expression profiles, *BMC Genomics*, 2006, **7**, 162.

24 R. D. Waite, A. Papakonstantinopoulou, E. Littler and M. A. Curtis, Transcriptome analysis of *Pseudomonas aeruginosa* growth: Comparison of gene expression in planktonic cultures and developing and mature biofilms, *J Bacteriol*, 2005, **187**, 6571-6576.

25 A. P. Morby, J. S. Turner, J. W. Huckle and N. J. Robinson, SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene smta - Identification of a Zn inhibited DNA-protein complex, *Nucleic Acids Res.*, 1993, **21**, 921-925.

26 R. B. Kapust, J. Tözser, J. D. Fox, D. E. Anderson, S. Cherry, T. D. Copeland and D. S. Waugh, Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency, *Protein Eng.*, 2001, **14**, 993-1000.

27 J. E. Tropea, S. Cherry and D. S. Waugh, Expression and purification of soluble His_6-tagged TEV protease, *Methods Mol. Biol.*, 2009, **498**, 297-307.

28 A. O. Pedersen and J. Jacobsen, Reactivity of the thiol-group in human and bovine albumin at pH 3-9, as measured by exchange with 2,2'-dithiodipyridine, *Eur. J. Biochem.*, 1980, **106**, 291-295.

29 D. W. Hasler, L. T. Jensen, O. Zerbe, D. R. Winge and M. Vašák, Effect of the two conserved prolines of human growth inhibitory factor (metallothionein-3) on its biological activity and structure fluctuation: comparison with a mutant protein, *Biochemistry*, 2000, **39**, 14567-14575.

30 T. J. M. Schoenmakers, G. J. Visser, G. Flik and A. P. R. Theuvenet, CHELATOR: An improved method for computing metal ion concentrations in physiological solutions, *Biotechniques*, 1992, **12**, 870-878.

31 O. I. Leszczyszyn, C. R. J. White and C. A. Blindauer, The isolated Cys_2His_2 site in E_C metallothionein mediates metal-specific protein folding, *Mol. Biosyst.*, 2010, **6**, 1592-1603.

32 H. Csopak, Specific binding of zinc(II) to alkaline phosphatase of *Escherichia coli*, *Eur. J. Biochem.*, 1969, **7**, 186-192.

33 C. Jacob, W. Maret and B. L. Vallee, Control of zinc transfer between thionein, metallothionein, and zinc proteins, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 3489-3494.

34 R. L. J. Keller, *Computer aided resonance assignment tutorial*, CANTINA Verlag, Goldau, 2004.

35 D. W. Byerly, C. A. McElroy and M. P. Foster, Mapping the surface of *Escherichia coli* peptide deformylase by NMR with organic solvents, *Protein Sci.*, 2002, **11**, 1850-1853.

36 P. Guerry and T. Herrmann, Comprehensive automation for NMR structure determination of proteins, *Methods Mol. Biol.*, 2012, **831**, 429-451.

37 P. Güntert and L. Buchner, Combined automated NOE assignment and structure calculation with CYANA, *J. Biomol. NMR*, 2015, **62**, 453-471.

38 Y. Shen and A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, *J. Biomol. NMR*, 2013, **56**, 227-241.

39 C. Baumann, A. Beil, S. Jurt, M. Niederwanger, O. Palacios, M. Capdevila, S. Atrian, R. Dallinger and O. Zerbe, Structural adaptation of a protein to increased metal stress: NMR structure of a marine snail metallothionein with an additional domain, *Angew. Chem. Int. Ed.*, 2017, **56**, 4617-4622.

40 M. E. Oates, P. Romero, T. Ishida, M. Ghalwash, M. J. Mizianty, B. Xue, Z. Dosztanyi, V. N. Uversky, Z. Obradovic, L. Kurgan, A. K. Dunker and J. Gough, D^2P^2: Database of disordered protein predictions, *Nucleic Acids Res.*, 2013, **41**, D508-D516.

41 R. L. Baldwin and G. D. Rose, Molten globules, entropy-driven conformational change and protein folding, *Curr. Opin. Struct. Biol.*, 2013, **23**, 4-10.

42 Z. Dosztanyi, B. Meszaros and I. Simon, ANCHOR: Web server for predicting protein binding regions in disordered proteins, *Bioinformatics*, 2009, **25**, 2745-2746.

43 B. Meszaros, I. Simon and Z. Dosztanyi, Prediction of protein binding regions in disordered proteins, *PloS Comput. Biol.*, 2009, **5**, e1000376.

44 V. N. Uversky, The most important thing is the tail: Multitudinous functionalities of intrinsically disordered protein termini, *FEBS Lett.*, 2013, **587**, 1891-1901.

45 B. A. Messerle, A. Schäffer, M. Vašák, J. H. R. Kägi and K. Wüthrich, Comparison of the solution conformations of human [Zn_7]-metallothionein-2 and [Cd_7]-metallothionein-2 using nuclear magnetic resonance spectroscopy, *J. Mol. Biol.*, 1992, **225**, 433-443.

46 G. R. Kowald, S. R. Stürzenbaum and C. A. Blindauer, Earthworm *Lumbricus rubellus* MT-2: Metal binding and protein folding of a true cadmium-MT, *Int. J. Mol. Sci.*, 2016, **17**, 65.

47 S. Zeitoun-Ghandour, J. M. Charnock, M. E. Hodson, O. I. Leszczyszyn, C. A. Blindauer and S. R. Stürzenbaum, The two *Caenorhabditis elegans* metallothioneins (CeMT-1 and CeMT-2) discriminate between essential zinc and toxic cadmium, *FEBS J.*, 2010, **277**, 2531-2542.

48 R. Dallinger, B. Berger, P. Hunziker and J. H. R. Kägi, Metallothionein in snail Cd and Cu metabolism, *Nature*, 1997, **388**, 237-238.

49 M. Höckner, K. Stefanon, A. de Vaufleury, F. Monteiro, S. Perez-Rafael, O. Palacios, M. Capdevila, S. Atrian and R. Dallinger, Physiological relevance and contribution to metal balance of specific and non-specific metallothionein isoforms in the garden snail, *Cantareus aspersus*, *BioMetals*, 2011, **24**, 1079-1092.

50 O. I. Leszczyszyn and C. A. Blindauer, Zinc transfer from the embryo-specific metallothionein E_C from wheat: a case study, *Phys. Chem. Chem. Phys.*, 2010, **12**, 13408-13418.

51 O. I. Leszczyszyn, H. T. Imam and C. A. Blindauer, Diversity and distribution of plant metallothioneins: a review of structure, properties and functions, *Metallomics*, 2013, **5**, 1146-1169.

52 O. I. Leszczyszyn, R. Schmid and C. A. Blindauer, Toward a property/function relationship for metallothioneins: Histidine coordination and unusual cluster composition in a zinc-metallothionein from plants, *Proteins Struct. Funct. Bioinform.*, 2007, **68**, 922-935.

53 O. I. Leszczyszyn, S. Zeitoun-Ghandour, S. R. Sturzenbaum and C. A. Blindauer, Tools for metal ion sorting: in vitro evidence for partitioning of zinc and cadmium in *C. elegans* metallothionein isoforms, *Chem. Commun.*, 2011, **47**, 448-450.

54 O. Palacios, A. Pagani, S. Perez-Rafael, M. Egg, M. Höckner, A. Brandstätter, M. Capdevila, S. Atrian and R. Dallinger, Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: Evolutionary differentiation of mollusc metallothioneins, *BMC Biol*, 2011, **9**, 4.

55 O. Palacios, S. Perez-Rafael, A. Pagani, R. Dallinger, S. Atrian and M. Capdevila, Cognate and noncognate metal ion coordination in metal-specific metallothioneins: The *Helix pomatia* system as a model, *J. Biol. Inorg. Chem.*, 2014, **19**, 923-935.

56 S. Perez-Rafael, F. Monteiro, R. Dallinger, S. Atrian, O. Palacios and M. Capdevila, *Cantareus aspersus* metallothionein metal binding abilities: The unspecific CaCd/CuMT isoform provides hints about the metal preference determinants in metallothioneins, *BBA-Proteins Proteom*, 2014, **1844**, 1694-1707.

57 E. A. Peroza, R. Schmucki, P. Güntert, E. Freisinger and O. Zerbe, The β_E-domain of the wheat E_c-1 metallothionein: A metal-binding domain with a distinctive structure, *J. Mol. Biol.*, 2009, **387**, 207-218.

58 A. W. Foster and N. J. Robinson, Promiscuity and preferences of metallothioneins: The cell rules, *BMC Biol*, 2011, **9**, 25.

59 C. A. Blindauer, N. C. Polfer, S. E. Keiper, M. D. Harrison, N. J. Robinson, P. R. R. Langridge-Smith and P. J. Sadler, Inert site in a protein zinc cluster: Isotope exchange by high resolution mass spectrometry, *J. Am. Chem. Soc.*, 2003, **125**, 3226-3227.

60 G. L. Öz, D. L. Pountney and I. M. Armitage, NMR spectroscopic studies of $I = {}^1/_2$ metal ions in biological systems, *Biochem. Cell Biol.*, 1998, **76**, 223-234.

61 L. Hemmingsen, L. Olsen, J. Antony and S. P. A. Sauer, First principle calculations of ^{113}Cd chemical shifts for proteins and model systems, *J. Biol. Inorg. Chem.*, 2004, **9**, 591-599.

62 J. Hidalgo, R. S. Chung, M. Penkowa and M. Vašák, Structure and function of vertebrate metallothioneins, *Met. Ions Life Sci.*, 2009, **5**, 279-317.

63 L. Vergani, Metallothioneins in aquatic organisms: Fish, crustaceans, molluscs, and echinoderms, *Met. Ions Life Sci.*, 2009, **5**, 199-237.

64 G. Platzer, M. Okon and L. P. McIntosh, pH-dependent random coil ^{1}H, ^{13}C, and ^{15}N chemical shifts of the ionizable amino acids: A guide for protein pK_a measurements, *J. Biomol. NMR*, 2014, **60**, 109-129.

65 D. S. Wishart, Interpreting protein chemical shift data, *Prog Nucl Mag Res Sp*, 2011, **58**, 62-87.

66 L. S. Busenlehner, N. J. Cosper, R. A. Scott, B. P. Rosen, M. D. Wong and D. P. Giedroc, Spectroscopic properties of the metalloregulatory Cd(II) and Pb(II) sites of *S. aureus* pI258 CadC, *Biochemistry*, 2001, **40**, 4426-4436.

67 K. H. Lee, C. Cabello, L. Hemmingsen, E. N. G. Marsh and V. L. Pecoraro, Using nonnatural amino acids to control metal-coordination number in three-stranded coiled coils, *Angew. Chem. Int. Ed.*, 2006, **45**, 2864-2868.

68 P. Faller, D. W. Hasler, O. Zerbe, S. Klauser, D. R. Winge and M. Vašák, Evidence for a dynamic structure of human neuronal growth inhibitory factor and for major rearrangements of its metal-thiolate clusters, *Biochemistry*, 1999, **38**, 10158-10167.

69 C. A. Blindauer and O. I. Leszczyszyn, Metallothioneins: unparalleled diversity in structures and functions for metal ion homeostasis and more, *Nat. Prod. Rep.*, 2010, **27**, 720-741.

70 M. Vašák, Dynamic metal-thiolate cluster structure of metallothioneins, *Environ. Health Persp.*, 1986, **65**, 193-197.

71 P. Gettins and J. E. Coleman, Cd-113 nuclear magnetic resonance of Cd(II) alkaline phosphatases, *J. Biol. Chem.*, 1983, **258**, 396-407.

72 J. D. Otvos and I. M. Armitage, Determination by Cd-113 nuclear magnetic resonance of the structural basis for metal ion dependent anticooperativity in alkaline phosphatase, *Biochemistry*, 1980, **19**, 4031-4043.

73 M. Vašák and R. Bogumil, Diversity of cluster structures in mammalian metallothionein: Interplay between metal ions and polypeptide chain, *Nato Asi 2*, 1997, **26**, 195-215.

74 S. S. Eaton and R. H. Holm, Inversion and ligand-exchange kinetics of tetrahedral bis-chelate complexes of zinc(II) and cadmium(II), *Inorg. Chem.*, 1971, **10**, 1446-1452.

75 K. S. Hagen, D. W. Stephan and R. H. Holm, Metal ion exchange reactions in cage molecules: The systems $[M_{4-n}M'_n(SC_6H_5)_{10}]^{2-}$ (M, M′ = Fe(II), Co(II), Zn(II), Cd(II)) with adamantane-like stereochemistry and the structure of $[Fe_4(SC_6H_5)_{10}]^{2-}$, *Inorg. Chem.*, 1982, **21**, 3928-3936.

76 N. Romero-Isart and M. Vašák, Advances in the structure and chemistry of metallothioneins, *J. Inorg. Biochem.*, 2002, **88**, 388-396.

77 C. A. Blindauer, Metallothioneins with unusual residues: Histidine as modulators of zinc affinity and reactivity, *J. Inorg. Biochem.*, 2008, **102**, 507-521.

78 T. Liu, M. Ryan, F. W. Dahlquist and O. H. Griffith, Determination of pK_a values of the histidine side chains of phosphatidylinositol-specific phospholipase C from *Bacillus cereus* by NMR spectroscopy and site-directed mutagenesis, *Protein Sci.*, 1997, **6**, 1937-1944.

79 B. A. Messerle, A. Schäffer, M. Vašák, J. H. R. Kägi and K. Wüthrich, Three-dimensional structure of human [113Cd7]metallothionein-2 in solution determined by nuclear magnetic resonance spectroscopy, *J. Mol. Biol.*, 1990, **214**, 765-779.

80 A. Arseniev, P. Schultze, E. Wörgötter, W. Braun, G. Wagner, M. Vašák, J. H. R. Kägi and K. Wüthrich, Three-dimensional structure of rabbit liver [Cd7]metallothionein-2a in aqueous solution determined by nuclear magnetic resonance, *J. Mol. Biol.*, 1988, **201**, 637-657.

Figure Legends

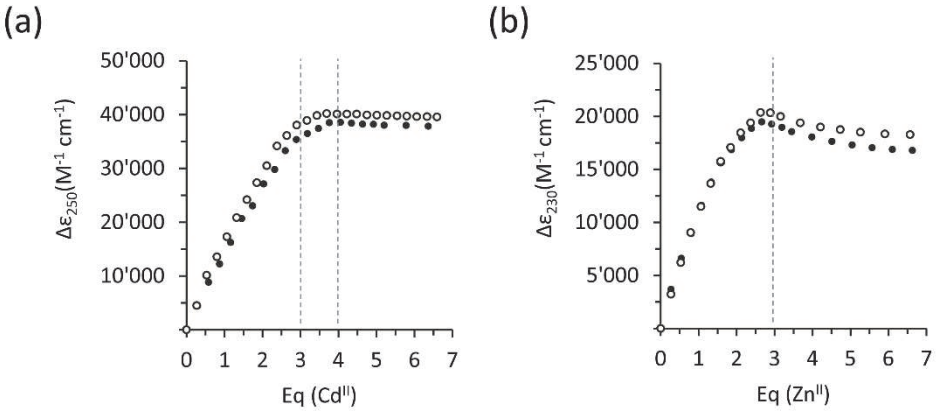


Figure 1. **Coordination of metal ions to full-length and shortened proteins.** Plot of LMCT band absorptivity at (a) 250 nm for the titration with Cd^{II} and at (b) 230 nm for the titration with Zn^{II} of apo-PflQ2 MT (◦) and apo-sh_PflQ2 MT (•).

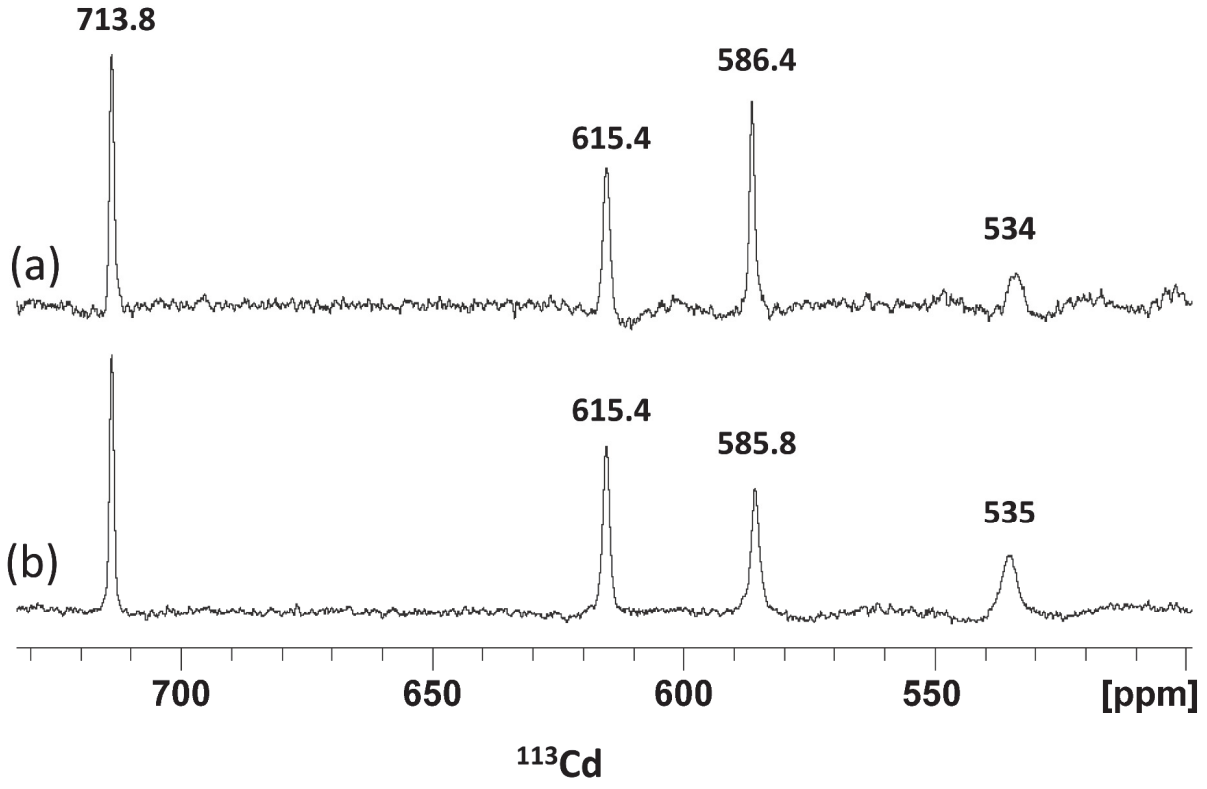


Figure 2. **^{113}Cd NMR.** Proton-decoupled 1D ^{113}Cd NMR spectra for (a) $^{113}Cd_4$sh_PflQ2 MT and (b) $^{113}Cd_4$PflQ2 MT at 320 K.

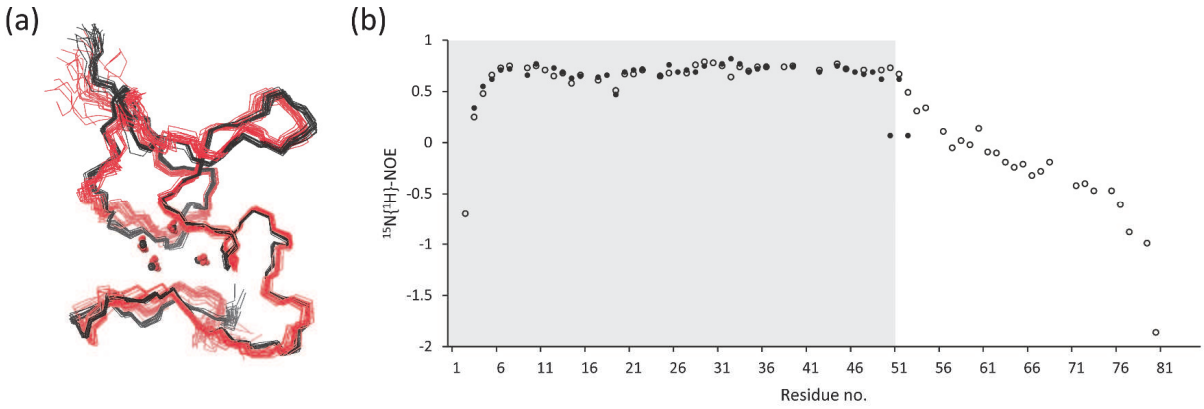


Figure 3. **Structure and dynamics of PflQ2 MT**. (a) Comparison of the protein backbone structures of residues 1-52 from Cd_4PflQ2 MT (black) and Cd_4sh_PflQ2 MT (red). (b) Comparison of $^{15}N\{^1H\}$-NOE values for PflQ2 MT (open circles) and its shortened version (full circles). The well folded domain is shaded in grey, Cd^{II} ions are shown as small spheres.

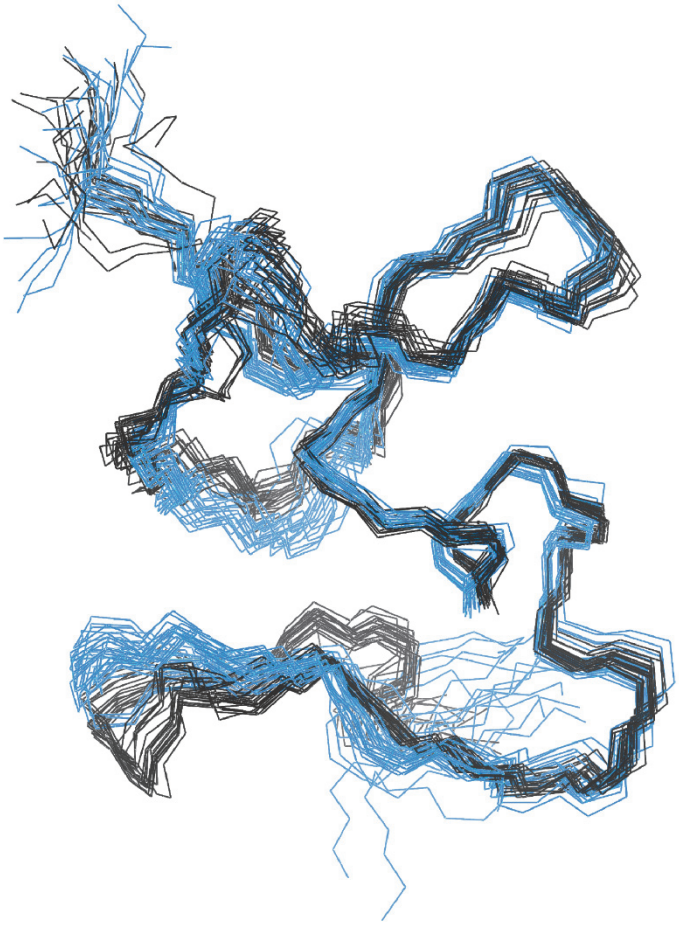


Figure 4. **Structures of the Zn^{II} and Cd^{II} species**. Comparison of the protein backbone of Zn_3- (blue) and Cd_4sh_PflQ2 MT (black).

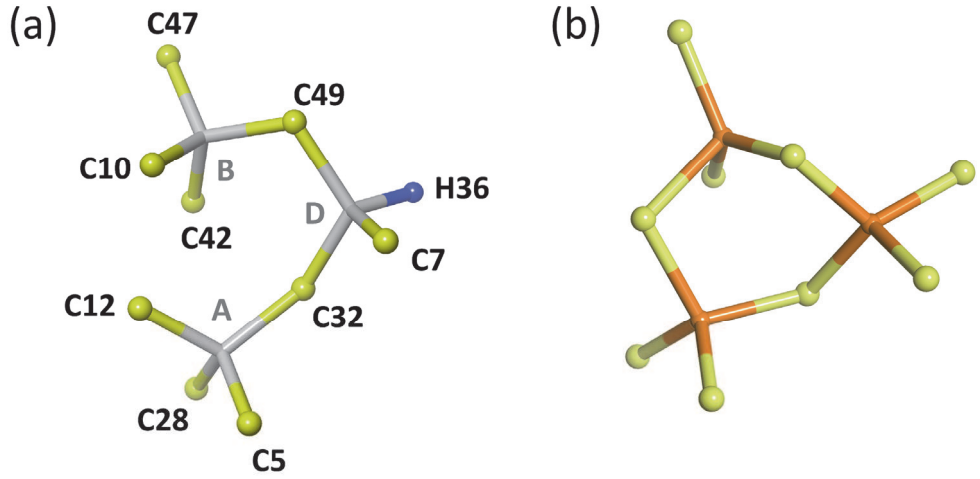


Figure 5. **Topology of the metal cluster in Zn_3PflQ2 MT** (a). For comparison the topology in the typical β-cluster of mammalian MTs is depicted in (b) (Cd_3Cys_9 β-cluster of human MT-2).[79]

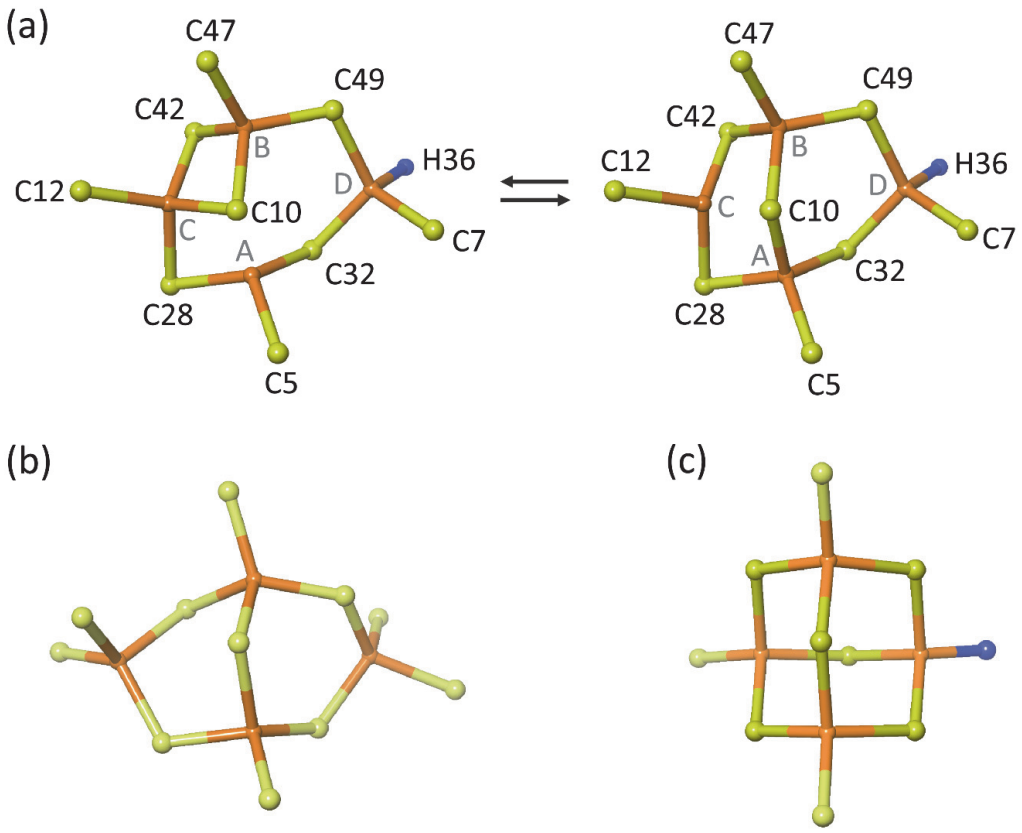


Figure 6. **Topology of the metal cluster in Cd_4PflQ2 MT.** (a) Interchanging arrangements of the cluster in Cd_4PflQ2 MT showing the two different bridging modes of residue Cys 10 (C10). (b) Typical α-cluster of mammalian MTs (Cd_4Cys_{11}-cluster of rat liver MT-2).[80] (c) Theoretical topology of a true-adamantane cluster with Cd_4Cys_9His stoichiometry.

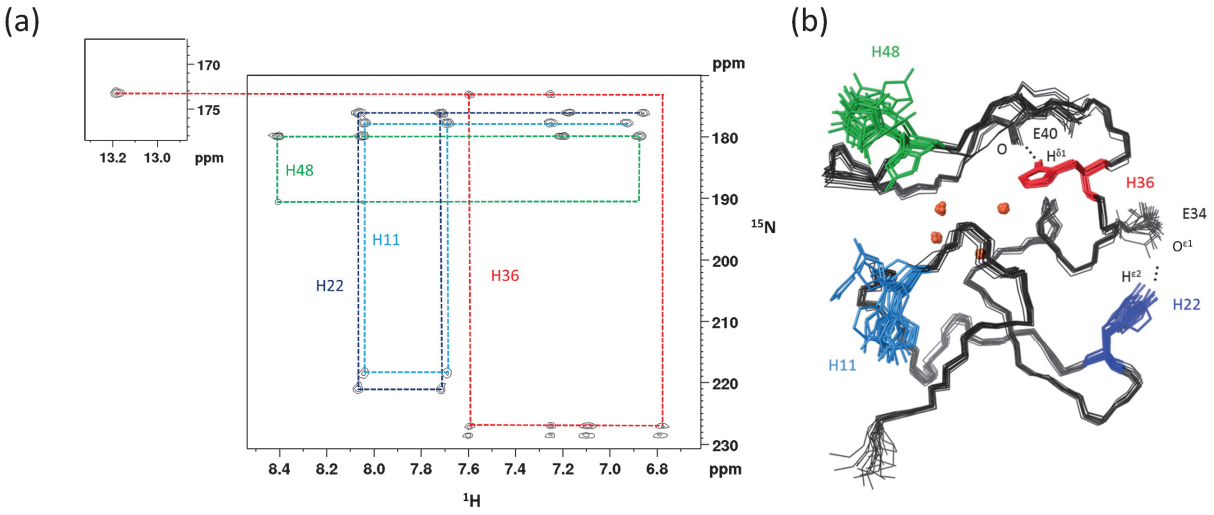


Figure 7. **Assignment and tautomeric states of His residues in Cd$_4$sh_PflQ2 MT.** (a) Assignment of tautomeric states of His residues in ^{13}C,^{15}N labelled ^{113}Cd$_4$sh_PflQ2 MT using a long-range [^{15}N,^{1}H]-HSQC spectrum (J(^{15}N, ^{1}H) =22.75 Hz; 310 K). (b) Ensemble of the 20 lowest-energy structures for Cd$_4$sh_PflQ2 MT with colour-coded His residues.

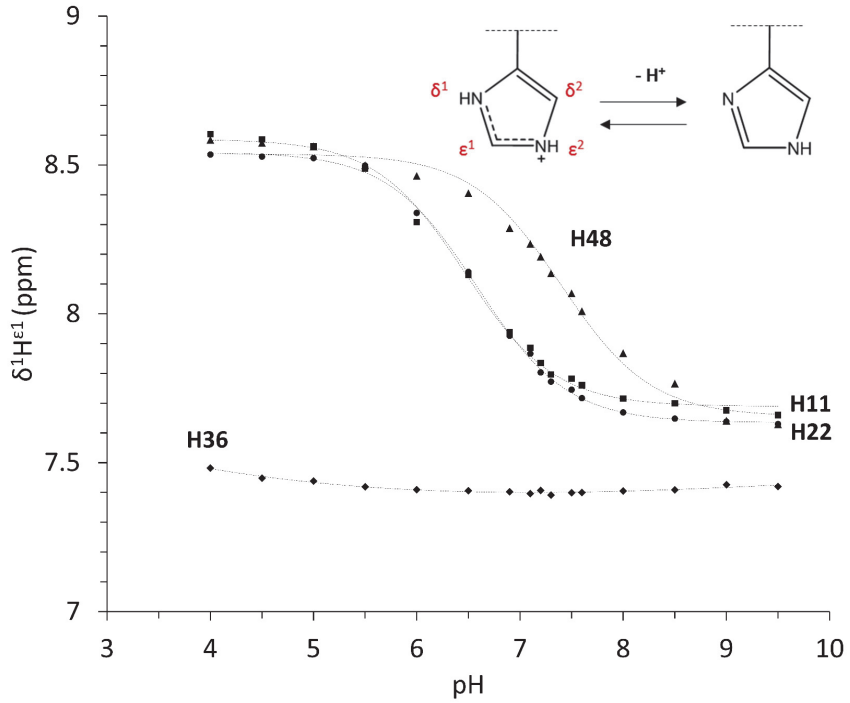


Figure 8. **Determination of His pK_a values by NMR.** pH dependence of H$^{\varepsilon1}$ chemical shifts of the four histidine residues in Cd$_4$sh_PflQ2 MT. Note that H36 cannot be deprotonated because of its involvement in Cd^{II} ion coordination.

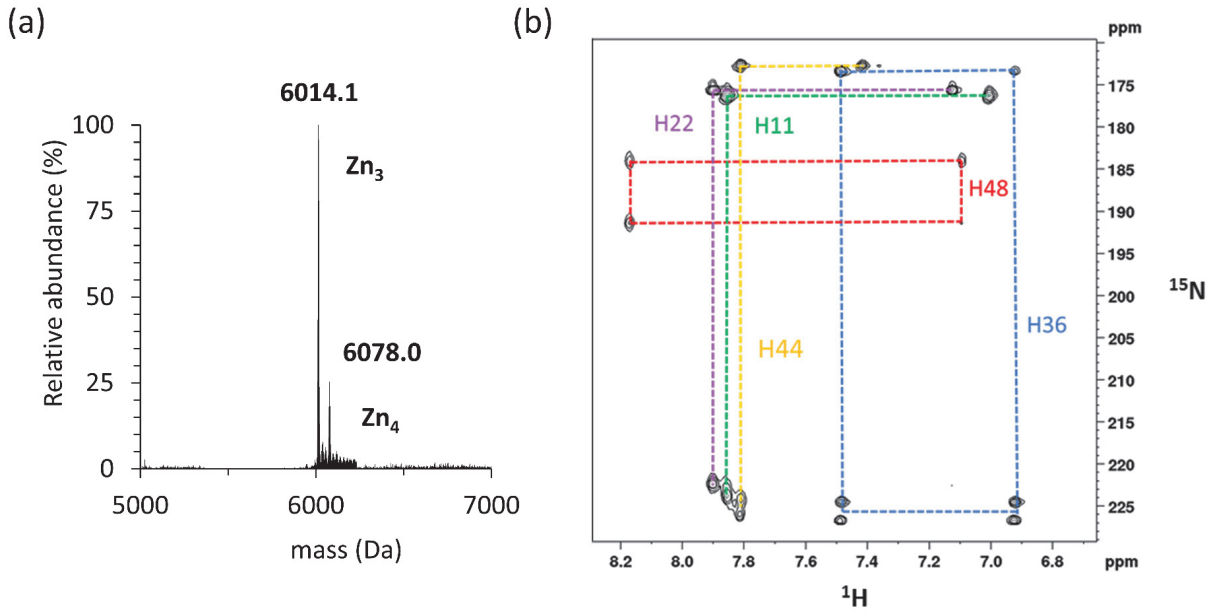


Figure 9. **Characterization of A44H-sh_PflQ2 by MS and NMR.** (a) Deconvoluted ESI(+) mass spectrum of Zn_4A44H-sh_PflQ2 MT (M_{calc} 6012.48 Da (Zn_3), 6077.86 Da (Zn_4)); (b) Long-range [^{15}N,^{1}H]-HSQC spectrum (J(^{15}N,^{1}H) = 25 Hz; 310 K) of ^{15}N labelled $^{113}Cd_4$A44H-sh_PflQ2 MT to determine the tautomeric states of the His residues.

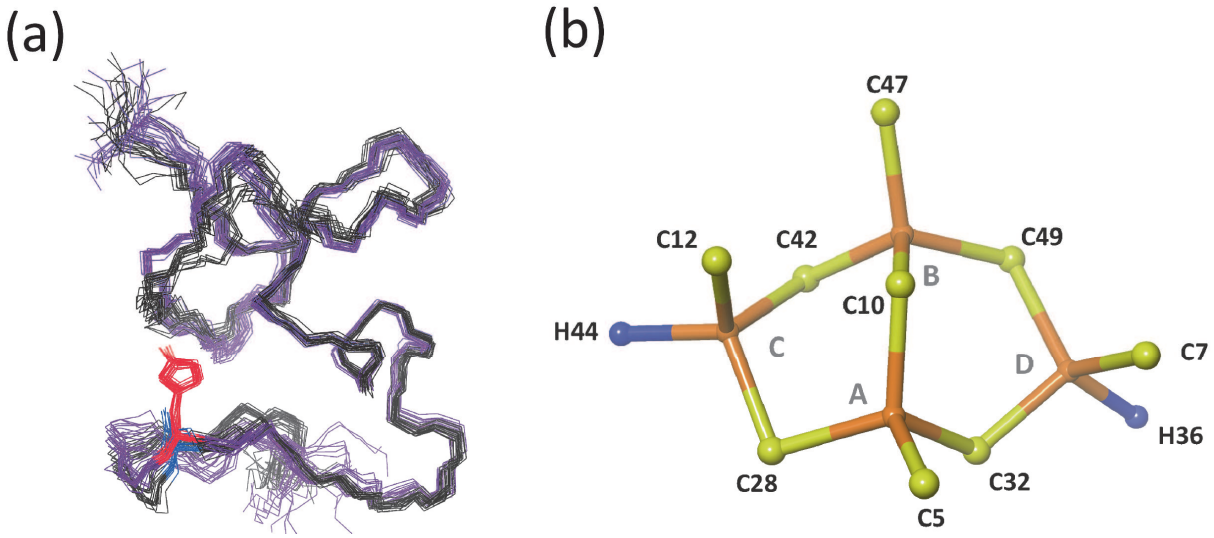


Figure 10. **Structure of Cd_4A44H-sh_PflQ2 MT**. (a) Comparison of the 20 lowest-energy structures of Cd_4A44H-sh_PflQ2 MT (purple) and Cd_4sh_PflQ2 MT (black). H44 is shown in red and A44 in blue. (b) Proposed topology of the Cd_4 cluster in A44H-sh_PflQ2 MT.